Minimality as a Universal Principle of Phonological Change

Marc Picard
Concordia University

Abstract

The minimality of phonological change is a tried-and-true principle that is meant to ensure that any proposed regular sound change will obey the all-important and fundamental criterion of naturalness by severely constraining the types of modifications that segments can undergo. It has shown itself to be an indispensable aid in setting up the most plausible and realistic reconstructions and relative chronologies by establishing clearcut demarcations between observed sound correspondences and actual sound changes. However, the validity and usefulness of this principle has recently been called into question by Scheer (2004) who argues that its application can lead to scenarios in which an inordinate number of steps can be proposed to link one historical segmental stage to another, or where no evidence, dialectal or otherwise, can be adduced to set up some intermediate stage that the principle of minimality would require to exist between two diachronically corresponding sounds. To this end, he adduces data involving (1) the velarization of dental stops in Cologne German, (2) the assibilation and fronting of velar stops in French, and (3) the Second German Consonant Shift. In this paper, I will review the cases described by Scheer in order to show that they do not contravene or
invalidates the principle of minimality in any way.

Keywords: sound change, universal constraints, minimality

1. Introduction

One of the primary constraints on any proposed regular sound change is that it must be minimal. This principle, which I have termed *the minimality of phonological change*, stems from an observation put forth by Donegan and Stampe to the effect that “processes represent responses to phonetic difficulties … and each process makes substitutions by altering a single phonetic property to remedy the difficulty. Since the substituted sound should, in each case, be as perceptually similar to the original target as possible, it follows that *the changes processes make will be minimal*” [my emphasis] (1979: 136-7). The same point has been made by other historical phonologists, notably by Bynon who states that since “changes in the realization of particular sounds must be small enough for speakers using both the old and the new realizations still to be able to recognize lexical items … it is usual for change to proceed in small steps which involve the alternation of only one feature at a time” (1977: 86).

This position has recently been challenged by Scheer (2004) who cites examples where the apparent application of minimality seems to lead to scenarios in which an unrealistic number of steps can be proposed to link one historical segmental stage to another, or where no evidence, dialectal or otherwise, can be adduced to set up some intermediate stage that the principle of minimality would require to exist between two diachronically corresponding sounds. In this paper, I will review the cases described by Scheer in order to show that although he is right to question the viability of the phonetically minute analyses he has brought to light, these do not contravene or
invalidate the principle of minimality that was presented in Picard (1990, 1994, 1999) which he cites.

2. Velarization in Cologne German

The first case study Scheer brings forth in order to show that “over-atomisation can lead to absurd results which escape any rational control” (2004: 75) involves the difference in the place of articulation of certain stops between Standard German (SG) and Cologne German (CG), as shown below:

(1)    Standard German       Cologne German
brown   braun   /braʊn/   brung   /bruŋ/
today   heute   /hɔjtə/   hück    /hʏk/
to cut  schneiden /ʃnaidən/ schnigge /ʃnɪgə/
colorful  bunt   /bʊnt/   bungk   /buŋk/
to bind  binden   /bɪndən/ bringe   /bɪŋə/

According to Scheer, the change from Middle High German (MHG) to CG is simply one whereby “dentals are velarised only if preceded by a MHG high vowel” (2004: 78), and he decries the fact that Heinrichs (1955, 1961) posits no fewer than six intermediate diachronic stages between dentals and velars. Thus, in order to account for the evolution of a form like MHG /tʃiitə/ (which is related to SG Zeit ‘time’) to CG /tʃiɡə/, Heinrichs proposes the following changes:

MHG      tʃiitə
Voicing   *tʃiidə
Deletion  *tʃiɪə
Schärfung (‘sharpening’) *tʃiiɪə
Anti-hiatus *tʃiɪjə
Scheer assumes that the principle of phonological minimality I subscribe to would require that I not only postulate this number of stages but a few more to boot given that “[a]ccording to this concept, a phonological process may only alter one property of a sound at a time” or, if “stated in terms of distinctive features, a process may modify only one feature at a time” (2004: 76). Thus, given this requirement, he concludes that “Heinrichs’ scenario falls foul of minimality because it shortens the root-vowel, inserts a hiatus-avoiding [j] and geminates this [j] in just one step” so that “the number of asterisked forms must be augmented to (at least) nine if minimality rules” (2003:76).

It is difficult to fathom how Scheer could have arrived at such a conclusion given that the principle of minimality I have formulated in a number of instances would not predict that any intermediate steps would need to occur between dentals and velars in the case at hand. As I have noted in previous studies (cf. Picard 1990: 86, 1994: 16, 1999: 68), the only requirement as regards minimality in the shift of one obstruent to another is that it can never simultaneously involve more that one of its three major phonetic properties, viz., voicing, place of articulation, and manner of articulation. Note, however, that the shift from obstruct to sonorant or vice versa can involve a simultaneous change in voicing and manner of articulation, as in Proto-Algonkian (PA) *l > Cheyenne /t/, e.g., *i1eniwa ‘man’ > /hiten/, or *l > Miami /l/, e.g., *a1emwa ‘dog’ > /a1emwa/. This is presumably because the feature “[+sonorant] is enhanced by [+voice] while [-sonorant] is enhanced by [-voice]” so that “the feature [voice] should take on the same value as the feature [sonorant] if the latter feature is to be implemented with maximum
strength” (Stevens and Keyser 1989: 89).

In sum, nothing that I have proposed anywhere in the past suggests that I would have dentals somehow inch their way from the alveolar ridge to the velum, or through some convoluted gradation such as that put forth by Heinrichs. My position in this matter is consistent with that of Blevins (2004) who cites a number of cases in various language families, and notably in four different groups of Austronesian languages, where /t/ has shifted to /k/, and for which she proposes no intermediate stage(s), viewing it simply as a shift from [coronal] to [dorsal].

Moreover, in my own study of Arapaho, a Western Algonkian language (Picard 1994), no transitional stage(s) such as /kp/, /kw/, /pW/, or the like were posited in the regular shift of PA *p to /k/, as in *papikwani > /kokúy/, *paašipahw- ‘to stab’ > /kooxúkohy-/. This was based on the fact that “the acoustic (and hence auditory) similarities between corresponding sounds made in the labial and velar regions is often considerable [because] sounds made with a constriction in either of these areas often have an overtone structure in which most of the acoustic energy is at a lower pitch than in the corresponding sounds made in the alveolar or palatal regions. The effect is greatest for voiceless stops, so that p and k sound more alike than p and t or k and t’” (Ladefoged 1971: 44).

3. Assibilation in French

The second case Scheer adduces in order to illustrate “[o]veratomised diachronic derivations that represent nothing but their author’s fantasy” (2004: 80) comes from Fouché’s (1961) analysis of the development of certain consonants and vowels from Latin to Modern French. The most extreme derivation Scheer cites is that of Latin /VkiV/ to French /Vs(V)/ as exemplified by the evolution of faciem /fakiem/ to face /fas(ə)/. The derivational history Fouché
presents is as follows (where superscripts and subscripts represent gradient phonetic stages such as postpalatal, mediopalatal, prelingual, etc.): $k > k^3 > k^{3,2} > kt_{2j} > kt_{s_{2j}} > tts_{s_{2j}} > ts_{s_{2j}} > ts > ts > s$.

Predictably, Scheer’s reaction to this sort of gradience is that “[p]honetic atomism of this kind, distinguishing no less than ten places of articulation between dentals and velars, is hardly appropriate for the phonological analysis of languages whose phonetic signal is available. Attempting to assign phonetic values of this precision and quantity to reconstructed forms can be no more than unexhausted ambition” (2004: 81). I could not agree more with this assessment since the principle of minimality I have always argued for and put into practice deals only with observable and well-attested phonological features and categories. This is closely akin to the view that has recently been expressed by Blevins:

First, phonological features are distinct from gradient phonetic properties, and typically reflect categories which have multiple distinct phonetic instantiations or cues. Second, no reference is made in phonological systems to non-contrastive features (e.g., the release of stops, 2 ms of voice-onset time, laminal interdental versus laminal dental, etc.). Each of these properties is integral to the explanations provided for typical cases of regular sound change. Because phonological features are distinct from gradient phonetic properties, and there is typically no one-to-one relationship between a phonological feature or category and a single phonetic characteristic, reinterpretations of the phonetic signal in the form of slight shifts of whole categories within the psycho-acoustic space are expected (2004: 90).

All this entails that in the passage of Latin /k/ to French /s/, only four intermediate stages would initially be proposed under the version of phonological minimality I have put forth:
PALATALIZATION (point of articulation)
\[ k > kj / __ [-cons, +front] \]

FRONTING (point of articulation)
\[ kj > tj \]

ASSIBILATION (manner of articulation)
\[ tj > tsj \]

DEPALATALIZATION (manner of articulation)
\[ tsj > ts \]

DEAFFRICATION (manner of articulation)
\[ ts > s \]

These processes would account for correspondences like the following:

(2)  kinkwe ‘five’  kentu ‘hundred’  bisak\textipa{k}ia ‘pouch’  merke˘dem ‘salary’
\[ sëk \quad sã \quad bëza \ s \quad mers \ i \]

Interestingly, this is exactly what has been proposed by Dauzat et al. (1971: xxii) and Pope (1934: 268), while other prominent analysts of French historical phonology have something very close to this, omitting only the affricated palatalized stage /tsj/ from their scenarios (cf. Brunot 1933: 163; ourciez 1967: 161, 171; Burciez & Bourciez 1967: 128-9).

In environments other than those described above, viz., word-initially and postconsonantally, the evolution of Latin /ki/ and /ke/ is slightly more complex. When these sequences are preceded by a vowel, they can end up as French /z/, e.g.,

(3)  plake:\textipa{r}e ‘to please’  wi:\textipa{k}i: ‘num ‘neighbor(ing)’  dikentem ‘saying’
\[ plëzi \ \textipa{R} \quad \textipa{vwazë} \quad dizã \]

However, this does not account for cases like that of \textipa{facciem} with
which we began this discussion, and in which Latin /Vki/ and /Vke/ do not undergo intervocalic voicing. Compare the following correspondences with those in (3):

(4) wikiam ‘vetch’ ma:kionem ‘mason’ wi:na:keam ‘marc’

The explanation that has generally been put forth to explain the difference between the two outcomes is based on whether Latin /Vki/ and /Vke/ were in hiatus or not, i.e., whether the sequences were /VkiV/ (> /VkjV/) and /VkeV/ (> /VkjV/), or /VkiC/ and /VkeC/.

As explained by Bourciez & Bourciez, preconsonantally “c (+ e, i) aboutit en français à [z], et il se dégage en avant un yod qui se combine avec la voyelle” (1967: 130), as shown in orthographic forms like plaisir (< placère) and voisin (<vīkīnum) in (3) above (this yod being absorbed by a high front vowel as in disant (< dicentem)). On the other hand, in forms like those in (4) where “[l]e yod ... procède de e, i latins en hiatus ... le résultat est ... le même qu’après une consonne”, that is to say, “il ne devient pas sonore, l’occlusive représentant c s’étant doublée en cours d’évolution” (131-2). In other words, “il est probable qu’en Gaule, dans tous les mots comme faciā, l’articulation de ky s’était renforcée de bonne heure en kky, d’où ensuite tts qui n’a pu devenir sonore ni au Midi ni au Nord (prov. fassa, fr. face)” (Bourciez 1967: 171). This is corroborated by Pope who states that “[w]hen kj is intervocalic the first element appears to have been lengthened and consequently escaped voicing” (1934: 130).

What stands out here is the fact that the analysis of the evolution of forms such as Latin faciam to French face that has been proposed by the most reputable French historical phonologists is in complete concord with that which is dictated by the principle of minimality. Though many intermediate stages are perforce involved in the derivation, all are natural, well-attested, and justifiable, respecting as
they do the criteria set forth above for shifts involving obstruents. Moreover, if one compares the number of diachronic stages involved in the derivation at hand with that proposed by Fouquet:

MINIMALITY (8 stages)
\[ k > kj > kkj > ttj > ttsj > tsj > ts > s \]

FOUQUET (10 stages)
\[ k > k^3 > k^{3.2} > kt_{2j} > kts_{2j} > tts_{2j} > ts_{2j} > ts_2 > ts > s \]

the latter does not appear quite as outlandish as Scheer makes it out to be. Finally, there is a serious flaw in his whole enterprise in that he fails to propose a scenario of his own. Is he suggesting that Latin /k/ became French /s/ in one fell swoop, and if not, exactly what steps would he omit? It is one thing to criticize but one should be disposed to present a plausible alternative solution.

4. The Second German Consonant Shift

Scheer’s most elaborate critique of minimality revolves around the following developments in the history of High German: “Unrecorded Common Germanic voiceless stops *[p, t, k] appear in Old High German as affricates [pf̩, ts̩, kχ] word-initially and after Codas (i.e., also in place of CG geminated voiceless stops), while fricative reflexes [f, s, χ/ç] are found in post-vocalic position (i.e., intervocally and word-finally after vowels)” (2004: 89). Following are the examples he provides:

(5) AFFRICATES
\[ pf:\ddot{a}:t \ ‘path’ \quad ts:\ddot{e}:n \ ‘ten’ \quad k\ddot{c}orn \ ‘corn’ \]

As noted by Scheer, “[t]he affrication of CG *k to [kχ] has occurred (or survived)
The only account of the distribution of these two types of obstruents that is consistent with naturalness in sound change is one which posits that all the voiceless stops in these various environments were first affricated, and that those affricates which were postvocalic were later reduced to fricatives. This is because a direct shift from [-continuant, +delayed release, -strident] segments like /p, t, k/ to [+continuant, -delayed release, +strident] segments like /f, s, χ/ is simply impossible. To falsify this, Scheer would have to find a sound change in progress, or at least of demonstrably recent vintage, where /p~f/, /t~s/, /k~χ/ are in perfect complementary distribution. Though many languages can be shown to have alternations like these, one would be hard put to find cases where these are the product of a transparent and phonetically conditioned synchronic rule.

If such alternations are impossible today, then they could not have emerged in the past since, as Lass has aptly noted, “the impossible never happened” (1986: 26). This stems from the two following uniformity principles:

A – PRINCIPLE OF PAN-TEMPORAL UNIFORMITY

Nothing (no event, sequence of events, constellation of properties, general law) that cannot for some good reason be the case in the present was ever true in the past.

---

only in High-Alemannic (Switzerland); the simple stop is found (or has been restored) elsewhere, thus [kɔrn] and [dæ̃kø] in N[ew]H[igh] G[erman]” (2004: 90).
B – PRINCIPLE OF UNIFORM PROBABILITIES

The distribution of probabilities in cases where a choice is available from a set of alternatives has always been in principle as it is now.

Thus, until such time as a sound change involving voiceless stops and (strident) fricatives can be found to be operating in some contemporary language, one should not be able to claim that such an event has ever occurred in the past.

As if this were not enough to disqualify his proposal, it turns out that this supposed stop-to-fricative transformation is even more unnatural since the output consists of geminates which are still reflected in the orthography of forms such as Wasser ‘water’ and offen ‘open’, a fact that Scheer does not deny as evinced by his statement that “[t]he real geminate value of the doubled consonants is beyond any doubt” (2004: 97). What he would have us believe, then, is that */p, t, k/ became /ff, ss, γγ/ in one fell swoop, a type of sound shift which remains universally unattested. In sum, he would rather live with the fact that “[t]he gemination of spirantised post-vocalic CG stops … remains mysterious” rather than have them “be derived from the twofold identity of affricates” (2004: 98) since he is convinced that affricates can occupy only one skeletal slot. However, since sequences like /pf/ and /ts/ which do not stem from any affrication process, as in English cupful and Betsy, can easily be conceived to undergo assimilatory changes to /ff/ and /ss/, it is difficult to see why /pf/ and /ts/ should not behave similarly since they are phonetically identical, i.e., [pf] and [ts].

In spite of all this, Scheer maintains that the affricates /pf, ts, ky/ and fricatives /f, s, χ/ that stem from the Germanic voiceless stops */p, t, k/ are the outcomes of two separate but simultaneous processes, i.e., that “both events are unitary and involve one single step” (2004: 91). More specifically, given that affricates are found
word initially and postconsonantally which, for some reason, he calls the Strong Position, while fricatives appear postvocalically – the so-called Weak Position – he arrives at the conclusion that

there was a general lenition process that attacked all CG voiceless stops, regardless of their position in the string. Consonants were more or less exposed to this lenition: they were relatively more protected in the Strong Position, whereas no such positional shield prevented them from experiencing full damage in weak positions. Hence, stops are fully lenited to fricatives in the latter case, but maintain an affricate status due to their positional protection in the former (2003: 91).

In sum, at some particular point in time, the following process was set in motion:

OLD HIGH GERMAN AFFRICATION AND FRICATIVIZATION

\[
p, t, k \rightarrow \begin{cases} ff, ss, \chi\chi \mid /C, #/ \mid V, Coda \end{cases}
\]

Here again, Scheer is invoking something that simply does not exist. In Picard (1994: 53-4), I examined such a possibility and came to the ineluctable conclusion that no empirical support exists for this type of sound change, thus warranting the postulation of the following universal constraint:

C — PRINCIPLE OF ONE-TO-ONE SEGMENTAL CHANGE
The simultaneous split of a segment X into two segments Y, Z is not a property of sound change.

In order to show that segments can actually go off into two different directions at the same point in time, he would have to
present instances where this has actually been observed to occur. I
know of no such cases.

At this point, one might ask what could have prompted Scheer to
replace such an eminently natural and widely accepted scenario as
generalized affrication followed by postvocalic fricativization with
one that seems so implausible, especially since he acknowledges (1)
that despite there having been numerous studies of these
developments since the 19th century, “nobody has adduced evidence
that refutes the possibility of a unitary affricate stage”, and (2) that
“[i]t is not reasonable to discard the solution that does not recognise
an intermediate affricate stage in weak positions for the simple sake
of being non-minimal” (2004: 99).

Scheer’s major and, in point of fact, only apparent argument
against the existence of a generalized affricate stage is summed up
in the title of one of his subsections: “There is not a shred of
evidence for post-vocalic affricates in any old or modern dialect”
(2004: 99). This is the conclusion he comes to after obtaining only
negative answers to questions like the following: “If affricates had
ever existed in weak positions, shouldn’t there be a trace in some
loanword or germanised place name? And, more importantly, is it
reasonable to buy into a scenario that wipes out every single
affricate in all dialects within a period of about three centuries?”
(2004: 100). His rationale, in sum, is that “grounding an argument
on the absence of a given form in the relevant dialectal area … may
enlighten diachronic events” (2004: 102).

Not only does Scheer’s whole line of argumentation fly in the
face of the principle of minimality but it also goes against
longstanding practice in historical phonology. One could cite
countless analyses of diachronic developments in all sorts of
recorded and unrecorded languages where unrecoverable
intermediate stages have been set up for various reasons. In fact, we
need look no further than the evolution of Latin /k/ to French /s/ we
examined in the last section for exemplification. In the evolution of
Latin *faciam* to French *fasse*, for example, very few of the intermediate stages /kj, kkj, ttj, ttsj, tsj, ts/ which have been postulated by the most reputable French historical phonologists are actually attested either historically or dialectally. This is particularly true of the hypothesized gemination (/kj/ > /kkj/) and subsequent degemination (/ttsj/ > /tsj/) whose only role is to prevent voicing, as evinced in the correspondence *placere:*plai*z*/ir. There is also no evidence of a palatalized /kj/ since all the Romance languages evince either an affricate or a fricative (or the original /k/ as in Sardinian) (cf. Bourciez 1967: 161).

5. Conclusion

In presenting the three cases we have examined above, Scheer’s objective was to attempt to discredit the principle of minimality of phonological change by showing that it must necessarily lead to a proliferation of unwarranted phonetic events in diachronic derivations or, at the very least, to some unnecessary and unproven intermediate stage(s) in such derivations. This was shown not to be the case in any of his examples. Thus, in the shift of /t/ to /k/ in Cologne German, minimality was shown not to lead to anything like the series of intermediate stages proposed by Heinrichs since such a development involves only point of articulation and is therefore permissible as a direct process. On the other hand, Latin /k/ to French /s/ in correspondences like *faciem:*face does require a number of intermediary steps if minimality is to be respected but this is exactly what the most prominent scholars have propounded without their having advocated or taken into account any such concept. Finally, Scheer’s refusal to acknowledge the existence of a postvocalic affricate stage between stops and fricatives in the history of German was shown to lead to unheard-of scenarios involving singleton stops spontaneously becoming fricative geminates and
simultaneously splitting into fricatives and affricates.

Overall, it is difficult to see where Scheer stands in terms of what should and should not be posited between sound correspondences. On the one hand, by stating that “[e]verybody agrees that two evolitional stages which involve a large phonetic distance are related by one or several intermediate forms” (2004: 101), he obviously does not subscribe to the what-you-see-is-what-you-get approach to historical phonology whereby a sound correspondence automatically equals a sound change. On the other hand, he fails to come up with any guidelines that would aid in determining what intermediate stages, if any, need to be posited in any given situation. In the end, he basically throws up his hands and claims that “the only thing on which we can ground our judgement is experience and intuition” (2004: 101), thereby propelling us back to the Dark Ages of diachronic phonological analysis where the most outlandish and unnatural processes could be boldly and impunibly advanced.

This free-for-all, anything-goes modus operandi is exactly what general conditions and constraints on sound change such as the principle of minimality were designed to counter and eliminate. There is a very limited number of directions in which individual sounds can evolve – what Catford refers to as “‘natural’ directions of development” (1974: 27) – and these are constant at any time or place. Thus, whenever one finds a diachronic sound correspondence between stops and fricatives, as in the case of German examined above, the question of whether there may or may not have been an intermediate affricate stage need not even be asked because it simply could not have been otherwise.

This does not mean that we can always determine the exact route that Segment A took on its way to becoming Segment B but we can at least determine that it could not have been A > B if this violates minimality. For example, as reported by Hyman (1975: 174-5), certain Bantu languages have the synchronic phonological rule p > s / __ i. By virtue of minimality, we would (in the absence of any
internal or external evidence of intermediate stages) be precluded from positing /p/ to /s/ as a direct diachronic change since it would involve both point and manner of articulation whereas Scheer would presumably have no such qualms, there being no principled reason for his rejecting it.²

In the final analysis, Scheer’s disgruntlement with minimality rests on the misconception that it is a phonetically based principle which inevitably leads to excesses in the postulation of intermediate stages between diachronically related segments. This is obvious from his categorical statement that “requiring diachronic change to be minimal is phonetic in essence, whereas allowing for non-minimal shifts is phonological” (2004: 102). Note, however, that at no time have I ever referred to the principle in question as anything but the minimality of phonological change. This is because, as mentioned above, it deals only with observable and well-attested phonological features which are known to “play a significant role in defining sound change and sound patterns” (Blevins 2004: 89), and which “are distinct from gradient phonetic properties” in that they “can reflect categories which have multiple distinct phonetic instantiations or cues” (Blevins 2004: 114). Under these conditions, there is simply no danger that derivations will ever fall victim to “over-atomisation” as a result of the operation of the minimality constraint.

² Though he does not reveal the nature of his evidence, Hyman lists the historical sequence of changes as pi > pʰi > pʰi > tʰi > si. He argues that “[e]ach of these historical changes is phonetically motivated” as “[t]he steps involved are (1) aspiration of obstructs before the high vowels /i/ and /u/, (2) affrication with an [s] release conditioned by the ‘grooved’ vowel [i], (3) assimilation of the place of articulation of the closure to the release of the affricate, and (4) deaffrication” (1975: 175). All of these changes are in conformity with the principle of minimality.
References


Minimality as a Universal Principle of Phonological Change