HOW MINIMAL IS PHONOLOGICAL CHANGE?

TOBIA S. SCHEER

1. Introduction

It is a very common situation that two recorded forms which contrast a diachronic relationship are not believed to represent two successive evolutionary stages. In regard to the fact that the phonetic distance between them is too important, no direct transformation involving one single phonological process is given credit. Rather, unattested intermediate forms are appealed to. Examples along these lines are countless; illustration may be given by the French palatalisation of Latin [k, g] before <e>: lat caru, gamba > Fr char, jambe [ʃar, ʒamb]. Ever since, French diachronians have proposed an intermediate unattested affricate stage *ʃ[ʃar, ʒamb], which is largely supported by Old French spelling <ch> (e.g. Bourciez 1926: 164ff., Pope 1934: 128, Fouché 1966-73 III: 555ff., Anglade 1949: 44, 72). This interpretation is also backed up by the numerous dialects that show present-day overt affricates as for example oriental Occitan near the Italian border: lat camba, cattu, capra, gau, gauta, gailina > [ʃambo, ʃat, ʃabru, ʃal, ʃawve, ʃalina] (Dalbera 1994: 422). Additional confirmation comes from loans that entered English at the time of the Norman conquest. These have preserved their affricate articulation up to now: engl channel, chief (Bourciez 1926: 165).

This particular case is prototypical for two reasons: it illustrates the common practice of assuming intermediate asterisked stages, which are corroborated by independent dialectal and/or migrational evidence. Since this pattern is frequent in diachronics, it has been credited of general value. Hence, no [s, z] or [ʃ, ʒ] may be viewed as the result of a single diachronic process whose origin is [k, g], even if there is no concomitant graphic or dialectal evidence. And even more so, this insight into phonological processes from diachronics has its synchronic mirror in the generative discussion of "naturalness" that was conducted in the 70s and early 80s. This debate has given birth to Natural Phonology (e.g. Stampe 1972, Dressler 1974) and Natural Generative Phonology (e.g. Hooper 1976, Vennemann 1974a, b). The question of abstractness in grammar that was raised by Kiparsky (1968) is closely related to this issue as well (e.g. Kiparsky 1982a, Tranel 1981). If the same phonological processes

---

1 The following pages are a reply to Marc Picard. They were inspired by his article "On spirantization and the minimalism of phonological change" (Picard 1999). I am therefore indebted to him in the first place. Discussion with Jean-Philippe Dalbera, Patrick Honeybone, Philippe Ségéal and Joaquim Brandão de Carvalho has further enriched my view on the matter.

2 Although some caution is in order here: more recent loans from French may show an affricate instead of the original fricative, e.g. gararge, which sounds [gzary] in French, but most commonly comes out as [garadʒ] in English.
Since the data are well known, I start the discussion directly with the state of affairs that is reconstructed for Common Germanic. The equations shown under (1) represent the three regular correspondences between Germanic and other IE languages.4

(1) 

<table>
<thead>
<tr>
<th>Common</th>
<th>IE</th>
</tr>
</thead>
<tbody>
<tr>
<td>recorded</td>
<td>Reihenschritt called</td>
</tr>
<tr>
<td>Germanic</td>
<td>Germanic (CG)</td>
</tr>
<tr>
<td>a. f.b.b, = *p.b,</td>
<td>*p.b, and *p^b,</td>
</tr>
<tr>
<td>b. b,d,g, = *B,D,G,</td>
<td>*B,D,G,</td>
</tr>
<tr>
<td>c. p,i,k, = *p,i,k, = b,d,g,</td>
<td></td>
</tr>
</tbody>
</table>

Tab. 1: The correspondence between Germanic and IE.

Based on this CG situation, recorded Germanic forms arose through the application of two processes: Verber’s Law (voiceless spirants become voiceless if the preceding vowel did not bear main stress in IE) and the strengthening of *B,D,G in contexts which are known to promote strength (Ségéral – Scheer 2001), i.e. roughly *B,D,G > b,d,g / (W,N) and in gemination. The implementation of the latter innovation shows great variation amongst Germanic dialects. Both processes have conjointly created the alternations between (voiced) stops and fricatives that are known as Grammatisch-Wechsel (e.g. OHG *žuhr → *džë:r̩ → *g < *de-dik’em en “tear 1st sg pres, 1st sg pret”).6

The picture under (1) is based on the following equations of recorded forms: (1a) got fadar, *hjutuha, *hafts = lat pater, tona, captus, (1b) got baieran, daur, OHG guas = gt fepo, gfe, zivui, (1c) got slupan, itan, *laulubur, *lubu, *lubus, etc. The detail of Grimm’s Law is exposed in, among many others, Schičed (1976), Hirt (1931: 79a), Klug (1913), Streiberg (1985), Prokoch (1939: 470), Penz (1975: 52f), Szulc (1974, 1987), Krahé (1969), Moventin (1949), Gnamelnitz (1964). Its classical presentation is cited here because it is needed to be amended if the glottalic theory (e.g. Gnamelnitz – Ivanov 1995, Collinge 1985: 63ff) is correct.

In the following table and henceforth, I use “B,D,G” for the series of Common Germanic voiceless spirants instead of the conventional stroke-out b,d,g.

4 The underlined letters are based on the fact that the voicing of fricatives (and not of voiceless stops) has been established by Paul (1874: 1906f). This view has become largely consensus, see Moventin (1954, 1972: 164f), Klug (1913: 49, 61), Streiberg (1985: 116f), Hirt (1931: 84f), Prokoch (1939: 330), Szulc (1974: 95), Krahé – Meid (1969: 90f, 97f) for further detail. It is supported by the fact that these reflexes of IE *p,t,k have undergone Verber and thus appear as voiceless consonants. However, the distribution of voiceless fricatives and stops of both origins is identical. As a consequence, the situation constitutes a common stage where IE *p,t,k have undergone Verber-modified IE *p,t,k and were merged and fricatives, i.e. CG *B,D,G, since the spirant character of the CG reflexes of IE *p,t,k is beyond any doubt.

5 Twadell (1939: 357), Prokoch (1918-19: 556), Hirt (1931: 84), Fouquet (1948: 27), Moventin (1954: 35f), Hammenr (1955: 26) and Penz (1975: 52f) for example are explicit on this argument.

An isolated view is expressed by Franck (1913) and Meillet (1908: 75, 89f; 1917: 45ff), who believe that Grimm’s Law has turned IE *b,d,g into voiced stops *b,d,g, which then underwent spirantisation into appropriate positions (i.e. intervocally). Franck and Meillet do not present any positive evidence in favour of IE *b,d,g > CG *b,d,g. They merely doubt the compelling character of the arguments that are put forth in support of IE *b,d,g > CG *B,D,G. However, the primitive character of *b,d,g in Germainic is refuted on grounds that have escaped Franck and Meillet (see Lessiak 1953: 272ff for a detailed response to Franck). For one thing, the aforementioned merger of IE *b,d,g and the Verber-allophones of IE *p,t,k is a fairly convincing argument. A related point in fact on which Paul (1874: 194), Lessiak (1933: 280) and Moventin (1954: 31f) draw attention is that Franck’s and Meillet’s scenario supposes a come-and-go (Lessiak: Ziecknick) between stops and fricatives that is certainly unwarranted and unnecessarily complicated: the “d” in a word like Gothic stadsis < IE *stati- “place-gen.” is a fricative. If Franck and Meillet are right, the derivation must involve the following four stages in

2. The genesis of the problem: Grimm’s Law

2.1. The facts and their received interpretation

Grimm’s Law stands at the outset of modern linguistics, comparatism and the neogrammarians’ claim that sound laws are as exceptionless as laws in natural science. The century-old debate on Grimm’s Law also illustrates the problem that is under scrutiny here: how many individual processes are represented by the facts in question?
2.3. Two events for three Reihenschritte: increased breath-pressure triggers aspiration and devoicing (neogrammarians)

Another attempt at reducing the three Reihenschritte of (1) to two events was undertaken on the assumption of a different causality: the increase of breath-pressure. However, two different Reihenschritte are supposed to be the consequence of a reinforced expiration: (1a) IE *p > pʰ > f and (1c) IE *b > p. As to the evolution of the former series, only the prelude IE *p > pʰ to (1a) is triggered by a stronger airflow. The actual spirantisation that results in CG *f belongs to a later stage. As may be seen, this scenario also assumes that IE *p has developed aspiration. However, this process is not interpreted as a spontaneous and hence causeless change. Rather, it stems from the alleged increase of subglottal pressure. The second consequence of this new articulatory habit is the "widening" of the glottis, which is provoked by the violent airflow and prevents the vocal cords from vibrating: IE *b > p is the consequence. In sum, (1a) and (1c) are understood as two different effects of the same fortition.

2.4. Two events for three Reihenschritte: aspiration and devoicing are one (structuralists)

The same division Tenues- (1a) and Medienverschiebung (1c) vs. Medialae Aspiratae-Verschiebung (1b) is operated by Kurylowicz (1948) and Fourquet (1948, 1954, 1955) (see also followers of the latter, e.g. Hammerich 1955 and Jung 1956), even though this is done for different reasons. Both analyses share most properties, but differ in one important aspect (neither author mentions the other).

Fourquet (1948: 10f) sets out to declare that what he takes to be the neogrammarians view is inaccurate: the three Reihenschritte at hand are not unrelated (in cause and time), nor do they stem from some (even unified) phonetic/physiological cause. Rather, the systemic pressure within the network of oppositions plays a central role. In other words, the minimal unit on which the diachronic action bears is not an isolated consonant or a series of consonants (a Reihe), but a correlation of oppositions: the evolution of IE *p > pʰ and IE *b > CG *p does not witness the transformation of two series, but of one single correlation: the original IE opposition *p: *b has become *pʰ: *p (*b: *p).* That is, the IE correlation of voicing has been abandoned in favour of a correlation of aspiration (Fourquet accepts spontaneous aspiration). Building on the typological recurrence of simultaneous p > pʰ and b > p (IE > Arabian, Modern Danish, High German Consonant Shift), Fourquet (1948: 47) claims that both innovations are concomitant and condition one another.

By this means, Fourquet hopes to replace a merely phonetic approach by a truly phonological account of Grimm's Law. However, his actual analysis is far from

---

8 Neogrammarians for the spontaneous pre-Germanic aspiration of IE *p,k,l are for example Paul (1874: 150ff), Kluge (1913: 51), Meyer-Benney (1901: 109). The former two authors consider a single cause (aspiration) for both spirantizations, i.e. *p (<IE *p) > CG *fand IE *b > CG *b. Lakoff (1968: 174) is a generative representative of the this view. The development of spontaneous aspiration has become commonplace since then; see, among others, Prokosch (1918-19:II-55, III-104), Leslau (1933- 280), Twaddell (1939: 339), Fourquet (1948: 47), Hammerich (1955: 21), Jung (1956: 297), Penzl (1975: 52). Synopsis views on the question are provided for example for Meyer-Benney (1901: 109f), Schröd (1976), Iversen – Salmons (1995).
9 For the sake of exposition, I use the labial series alone in order to represent the four places of articulation. Everything that is said about dentals and (labio-) velars as well.

meeting his structuralist ambitions. Fourquet by and large rests on the same phonetic grounds that were invoked by his neogrammarian predecessors.

The one difference between Fourquet's (1948) and Kuryłowicz's (1948) scenario is that the former author merely observes the devoicing of IE *h, while the latter provides a truly systemic reason for that fact: voicing has lost its function in the new system since the old correlation of voicing (IE *p vs. *b, *bh) has been replaced by a correlation of aspiration (pʰ, bʰ vs. *b). Features without systemic function are redundant and therefore evacuated: IE *h devices because IE *p has become aspirated (Hammerschich 1955: 21 also refines Fourquet's analysis in this sense).

In sum, the structuralist approach reduces the three Reihenschrifte to two events, dividing (1b) IE *p > CG *f and (1c) IE *h > CG *p on one side, against (1b) IE *bh > CG *B on the other. This partition is identical with the one that was operated by those who believe in increased breath-pressure.

2.5. One event for three Reihenschrifte: increased breath-pressure causes everything

Finally, the drive towards a unification of the Reihenschrifte has also produced analyses that aim at identifying one single cause for all three events. This is the case of Meyer-Benfey (1901) and Zabrocki (1951). The former author extends the effects of the alleged particularly intense breath-pressure to (1b) IE *bh > CG *B. On his account, aspiration (IE *p > *pʰ) and devoicing (IE *h > CG *p) are not the only result of an intense airflow. Both spirantisations (i.e. *pʰ > CG *f and IE *bh > CG *B) are also due to this cause. All other authors refrain from registering spirantisation as some kind of fortition for obvious reasons: spirantisation goes down the sonority scale and thus, if anything, falls into the category of lenition. Karsten (1939: 116) for example is explicit on this obstacle against an overall unification. According to Meyer-Benfey (1901), however, aspiration and spirantisation of IE *p are gradual and successive consequences of an increased airflow, which ultimately causes the loss of occlusion.

Zabrocki (1951) explicitly identifies the overall goal of his study: he wants to overcome the "neogrammarian phonetic atomism". What this means is that the events observed should not be viewed as various phonetic processes that are unrelated in time and causality. Rather, a common cause for the three Reihenschrifte should be sought. The same charge has been raised against neogrammarians in the structuralist literature. Fourquet (1948: 10, 17) for example rejects the "atomism of the diachronic method" (see also Szulc 1974: 105ff, 1987: 65ff).

2.6. Summary: a collective and century-old effort to reduce the number of processes

As was stated in the introduction, classical reconstruction supposes intermediate asterisked stages between two recorded forms whose phonetic distance is judged too important. This tradition has been transformed into a rigid principle by Picard (1990, 1994, 1999) who allows any non-minimal diachronic change. Now the study of Grimm's Law since the 19th century seems to indicate a reverse movement: instead of multiplying intermediate stages, research has sought to reduce three equations that were formerly held to be independent in cause and time to two phonological processes (or even one). The work in the post-war structuralist tradition clearly identifies the unwarranted feature of former analysis: neogrammarian phonetic and diachronic atomism. Unnecessary microscopic intermediate steps veil the actual phonological process and must be overcome.

Both approaches are well grounded but appear to be contradictory. The purpose of the remaining pages is to evaluate whether one or the other should be given credit, or whether a non-dogmatic case-by-case practice can be recommended.

3. Minimal excesses

This section examines whether Picard's minimal request can serve as an absolute measure for the quality of diachronic analysis: does a scenario that invokes x⁺1 intermediate steps guarantee a closer match of reality than an alternative with only x intermediate forms? Two case studies show that over-atomisation can lead to absurd results which escape any rational control.

3.1. Heinrichs' analysis of Cologne German velarisation

In order to introduce the matter, let us consider the treatment of velarised forms in Cologne German (henceforth CoG) by Heinrichs (1955, 1961). In the dialect spoken in the city of Cologne, words that possess the velars [t,d,n,nt,nd] in other varieties (among which Standard German, i.e. New High German, NHG) occur with velars. The following table provides illustration.11

<table>
<thead>
<tr>
<th>(2)</th>
<th>NHG</th>
<th>CoG velar</th>
<th>NHG</th>
<th>CoG</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>n</td>
<td>braun</td>
<td>[braun]</td>
<td>[braun]</td>
<td>brown</td>
</tr>
<tr>
<td>t</td>
<td>k</td>
<td>heute</td>
<td>[hteat]</td>
<td>[htek]</td>
<td>today</td>
</tr>
<tr>
<td>d</td>
<td>g</td>
<td>schneiden</td>
<td>[snaiden]</td>
<td>[snik]</td>
<td>to cut</td>
</tr>
<tr>
<td>b</td>
<td>nt</td>
<td>bunt</td>
<td>[bunt]</td>
<td>[bunt]</td>
<td>to bind</td>
</tr>
<tr>
<td>nd</td>
<td>g</td>
<td>binden</td>
<td>[bden]</td>
<td>[bene]</td>
<td>to bind</td>
</tr>
</tbody>
</table>

Supposing that Middle High German (MHG) and Old High German (OHG) are the ancestors of CoG, we may call on three (and only three) recorded stages for the purpose of diachronic derivation: CoG, MHG and OHG. Taking as an example OHG=MHG zit [ziit] = NHG Zeit [ziit] = engl. tide = dutch tijd = swedish tid, Heinrichs (1955: 240ss, 1961: 104) proposes the following intermediate asterisked stages for those inflected forms in which the root-final consonant appears

---

11 Classical descriptions of the CoG situation may be found in Münch (1904), Stark (1970), Frings-Schmitt (1942), Bertram (1935), Müller (1942), Kuepper (1992); a complete record of velarised items is available in Scheher (ms). The situation observed in Cologne is a particular case of what is known as the Rheinisch Velarisation, which in fact extends to a much broader geographical area, see for example Werlen (1983), Frings (1956).
The sound changes he proposes make reference to four types of motivation: 1) attested events in neighbouring dialects, 2) in other languages, facts of 3) general phonetic and 4) general phonological nature. The detail is as follows.

Loss of [d] in intervocalic position is observed in an adjacent area to the North and West of Cologne in such words as [soan], [rus] = NHG roten "red art", cf. Aubin et al. (1926: 125f). There is no loss of intervocal [d] in CoG, but Heinrichs points out the fact that this change also occurs in Luxemburgish, which is far South of Cologne and suggests that the area of d-loss once extended farther to the South than it does today, and included Cologne.

"Schrürung" is a term which is traditionally used to characterise the (allegedly) peculiar Rhenish accent. In particular, it is supposed to bear influence on vowel quantity, cf. Heike (1962). This part of Heinrichs' derivation is not necessary in order to relate the two attested forms. It is probably present because its occurrence is assumed for all Rhenish dialects.

Heinrichs does not provide any special motivation for the widely attested process that inserts a consonant in order to avoid a hiatus. The fact that in this case, the epenthetic consonant is [j] rather than anything else is not motivated either.

For the velarisation itself, Heinrichs invokes the evolution from Common Germanic *[jj] to Old Nordic [gg] (Holzmans Law) as in CG *twajjo > ON tveggja (=got twajjo) (e.g. Braune 1884, Collinge 1985: 93ff). In addition, he proposes an intermediate fricative stage *[yy].

Heinrichs also assumes identical behaviour of [d] in n_v and v_v. He motivates this equation by the phonetic and phonological identity of [n] as a sonorant. Sonorants are the most vocical consonants, and they often pattern with semi-vowels or vowels. As a witness, Heinrichs calls on the organisation of Germanic strong verbs into seven ablauting classes. Classes six and seven which are not involved in qualitative ablaut laid aside, the remaining five classes are defined by the right context of the root-vowel (1._j, 2._w, 3._NC, 4._licU, 5._C). The first four classes, but not the fifth, may be said to be "diphthongal" according to Heinrichs, to the effect that nasals and liquids behave like second parts of diphthongs.

Finally, Heinrichs anchors the alternative change VdV > VVV in the observation of parallel events in Danish and other German dialects.

In sum, Heinrichs' scenario is based on the idea that the intervocalic loss of [d] is at the origin of velarisation. This is one reason for which he has to invoke analogy with inflected forms: MHG NAsg zlt do not place the root-final consonant in intervocal position. In his opinion, support for the analogical solution also comes from his own dialect spoken in Amern, Kreis Kempen-Krofeld, where [d] is lost in intervocal position, with a resulting opposition NAsg [tit] vs. Npl etc. [tijja]. Amern, then, has not carried out levelling, as opposed to CoG where the [t] of NAsg has become [k] under the influence of the other inflected forms.

3.2. The arguments on which asterisked forms may be grounded

The purpose of the foregoing discussion is not to evaluate how close Heinrichs' diachronic derivation matches reality. There is little chance that it corresponds to
anything that has ever been pronounced since there is no loss of intervocalic [d] in Cologne at all. And supposing that such a process has been active in former times does not explain where the actual intervocalic [d]s that occur in CoG and faithfully continue the old state of affairs come from. More dramatically, Heinrichs' account does not even mention the critical factor that conditions CoG velarisation: dental velarisation only if preceded by a MHG high vowel, a fact which has been evidenced a century ago, cf. Höning (1877: 18ff), Münch (1904: 42, 97).

Be that as it may, it is interesting to observe on which kind of arguments Heinrichs grounds his reconstruction. His sympathy for a given asterisked form rather than for another is guided by four patterns.

First, an intermediate form or process may be appealing because it is attested in a neighbouring dialect. This neighbouring state of affairs may then be interpreted as an intermediate diachronic stage that the dialect under interest has gone through. In our example, this pattern is illustrated by the loss of intervocalic [d] and the opposition between Nasg and the other forms.

Second, motivation for a suggested asterisked form or process may be sought in events that occur in languages which do not share any geographic and/or genetic kinship with the dialect at stake (in the relevant period). Heinrichs uses this kind of reference in the Old Nordic case, as well as when he invokes Danish and other German dialects for Vd V > VOV. The same kind of argument is implicitly made when he suggests that consonant epenthesis is perfectly plausible in hiatus position because cross-linguistic observation provides countless examples of similar hiatus resolution. This type of evidence may be called typological.

Finally, general phonetic and phonological properties of segments may be used to assess intermediate forms. This is what Heinrichs does in his discussion of [n] as a sonorant.

In order to complete the list, three more arguments need to be mentioned. On one hand, of course, examination of the morpho-phonological variation within the language considered can provide precious clues. This evidence is commonly referred to as internal reconstruction. On the other hand, structuralist analysis has introduced two additional arguments: systemic pressure and functionalist reasoning. The former holds that a process may take place in order to rebalance a system. That is, either a systemic "hole" may be filled through "aspiration" ("drag-chain"), or a systemic overload for a given place/manner etc. may provoke dissimilation ("push-chain"). The latter states that an innovation may be triggered by the fact that a given sound has lost its (distinctive, systemic, communicative etc.) function, or acquired a new one (grammaticalisation). One case in point is Kuryłowicz's (1948) treatment of IE *b > CG *p that was discussed earlier.

More discussion of the arguments that are used in order to determine intermediate asterisked (or primitive) forms, as well as of the phonetic reality that they are supposed to represent, is provided for example by Lass (1997: 216ff), Campbell (1998: 117ff), Trask (1996: 231ff), Hock (1991: 535ff, 568ff). The following table summarises the arguments that usually lead to the positing of intermediate asterisked stages.

(4) a. dialectology: evidence from neighbouring languages that are genetically akin in the relevant period.
   b. typology: e.g. "the insertion of an anti-hiatus consonant within two independent vowels is a frequently observed process in many genetically unrelated languages".
   c. phonetic and phonological universals: e.g. "sonorants behave like vowels because of their inherent sonorant properties".
   d. systemic pressure: systemic aspiration or overload factor: an object that has been deprived of, or has acquired a function (distinctive, communicative, systemic etc.)
   e. internal reconstruction

3.3. How do we evaluate concurrent scenarios that are not implausible?

But what do these arguments tell us? They certainly allow to discredit certain asterisked forms. For example, a putative insertion of a vowel in order to break up a hiatus will not be taken seriously because cross-linguistic observation shows that this process does not occur in natural language. A famous illustration of this line of thought is the glottalic theory developed by (among others) Gamkrelidze and Ivanov (1995), which has led Indo-europeanists to accept the control of typological data over reconstructed forms. The classical Brugmannian system of IE consonants could not explain the striking infrequency of *b. Very few roots seemed to occur with that sound, and most of them were of onomatopoeic nature, e.g. Szemerényi (1990: 159f). There are synchronically accessible languages such as many Semitic languages including Arabic that lack a labial plosive, but in those languages [p] is missing, not [b]. In systems with ejective plosives such as those that occur in Kartvelian languages on the other hand, [p] is poorly attested, if not absent. Based on this synchronic typological observation, Gamkrelidze and Ivanov (1995) conclude that the real phonetic value of the traditional IE voiced plosives is 1) voiceless and 2) ejective (= glottalic).

Hence, a proposal may be abandoned if it falls foul of general typological data. The same holds true for the other arguments mentioned. But what is their incidence on concurrent asterisked forms that are not implausible from an internal, dialectological, typological, phonetic, phonological, systemic and functionalist point of view? They do not appear to be of any help in this kind of situation.

In the case of the CoG velarisation, at least two competing proposals have been advocated. Bruch (1954) argues for "Verhärtung" (hardening), i.e. MHG zit > [zit] > *[zi] > *[zi] > [zi] > [z] involving no analogy, no loss of intervocalic [d] and a [c]-stage preceding immediately velarisation. Martin (1922: 52) presents yet another way of relating MHG zit to CoG zik: on his account, [t] went first to [z], which then evolved to [k], i.e. MHG zit > [zi] > [zik]. He could quote British English glottalisation in words such as city [siti], sit [sit] etc. in support of his solution.

Is there a way to assert that any of the three reconstructions matches reality any closer than the others? I do not believe there is, at least not on the grounds of the arguments discussed. We could try to make a point on the basis of phonetic statistics: "the sounds X and Y do both exist, but X is five times more frequent than Y in the world's languages", or "both processes X and Y exist, but X is seven times more frequent than Y". Yet, we would not get any handle on reality if we gave systematic precedence to X in these cases, for the simple reason that in spite of X's frequency, Y does exist. On a
hundred reconstructions, we would reconstruct one hundred times X, and dismiss the fact that Y should occur twenty times if it is four times more infrequent than X.

In this situation, Picard (1999) argues, the minimal criterion can help out: non-minimal derivations are disqualified, and the minimal crown identifies the best solution. Before discussing Picard’s proposal in further detail, let us review another case of over-atomisation, which illustrates common practice in the traditional landscape.

3.4. Over-atomisation: Gallo-Romance viewed by Pierre Fouché

Over-atomised diachronic derivations that represent nothing but their author’s fantasy are hard to disqualify on rational grounds. One case in point is Fouché’s (1966-73) way of deriving Modern and Old French forms from Latin. Consider the following derivations that show his path for lat ki, #ke and tonic short e, o.

13 Table (5) reads as follows. I have translated Fouché’s (1966-73) idio-syncretic symbols into IPA wherever I could. Nonetheless, “t, k, t’, k’” remain unidentified, since Fouché does not provide any phonetic description (“E” is my own place-holder for something that Fouché notes as “x”, but which is different from [ε]). An acute accent on a vowel denotes stress. The chronology of the changes which have affected the sounds that are not at stake here has been neglected. For example, Latin final [a] is schwa in Old French, but there is no point in trying to single out one of the seven asterisked stages of the word *facia where this vowel reduction could be “observed” for the first time. Fouché reverts to a sophisticated system of superscripts and indexes in order to transcribe phonetic detail, but is not consistent at all in their use throughout the book. I reproduce hereafter Fouché’s (1966-73 I:41) glosses for the various indexes and superscripts concerning dental, alveolar, palatal and velar.

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>apico-coronales</td>
<td>pré-dorsales 1</td>
<td>médio-dorsales 2</td>
<td>post-dorsales</td>
</tr>
<tr>
<td>dentales alvéolaires</td>
<td>nivo-dépalliales 1</td>
<td>prépa-médio-dépalliales 2</td>
<td>postpa-médio-dépalliales 3</td>
</tr>
<tr>
<td>t, t’</td>
<td>f</td>
<td>s</td>
<td>s</td>
</tr>
</tbody>
</table>

Hence, given the boldfaced numbers, C’s describes a consonant whose production involves x as a passive, and y as an active articulator; y describes the part of the tongue (first line under the Roman numbers) that touches or approaches the portion x of the palate, velum etc. (second line under the Roman numbers). If only one superscript is mentioned, it refers to the tongue. Fouché (1966-73:41) says that most of the sounds described by his table do not exist anymore in Modern French, as for example the voiceless plosives [t, t’, k, k’]. Consequently, one may assume that [t, k, k’] still exist and are distinct from one another, a fact that demonstrates Fouché’s rather acute sense of observation. The destines of *facia, centu, pede and mola are quoted from Fouché (1966-73 III: 911, III: 553f, IV: 219f, 265f, IV: 204f), respectively.

14 Much in the way of “phonetic” characterisations such as “increase of breath-pressure” and the like, on which neogrammarians and structuralists ground their analyses, cf. section 2.
minimalism can certainly not be taken as an absolute certificate of quality: Fouché demonstrates that less minimalism may in some cases come closer to reality.

4. What does minimality buy us?

4.1. Do we want to know what happened in the mouth or in the brain?

Let us now return to the case of several competing diachronic scenarios that are judged equally plausible on the grounds of the criteria under (4). Picard (1994, 1999) suggests that different proposals in this situation may be ranked according to a universal principle that makes exclusive reference to their formal properties: non-minimal derivations must be rejected. Picard’s understanding of “minimal” in this context refers to distinctive feature theory: a sound change from X to Y is minimal iff the difference between X and Y corresponds to the contrast of the (binary) value of exactly one feature. Changes that involve two articulatory moves, say, dental > alveolar > palatal, are accepted as minimal if the feature system used does not reflect the difference between alveolar and dental/ palatal. If on the other hand the particular feature geometry chosen does express a three-way distinction in this area, the change will be outlawed, and an intermediate step required. On this account, a change [k] > [l] such as in our initial French example cannot be minimal because stopness and the place of articulation are controlled by two distinct features in all feature geometric systems. Including an intermediate stage *[f] makes the derivation minimal because only one feature needs to be manipulated for every step.15

This definition of minimality and hence the degree of asterisked atomism that is required crucially relies on a theoretical prerequisite, i.e. the particular feature-geometric theory that is used. More generally speaking, the fineness-of-grid depends on the particular notational system that legislates: “we can only see what our notation allows us to write” (Lass 1997: 221). Since Jakobson’s acoustic and SPE’s articulatory approach, debate in the area of distinctive feature systems has concerned the number, the label, the arboREAL organisation and the valency of features, as well as their possible underspecification. The discussion is far from being conclusive, and a consensual model has not yet emerged (see for example Clements – Hume 1995, Steriade 1995). Moreover, there is another approach to the internal organisation of speech sounds which roots in the recognition of primitive melodic primes (also called monovalent, categorical, holistic), rather than of binary features. This line of thought was introduced by Anderson – Jones (1974); it has been implemented in Dependency Phonology (Anderson – Jones 1977, Anderson – Ewen 1987), Government Phonology (Kaye et al. 1985) and Particle Phonology (Schane 1984). Since contrast is not expressed by different values of a given feature but by the presence vs. the absence of a prime, entirely different predictions are made as to what counts as minimal. Needless to say, if the debate on the precise properties of feature geometry is ongoing, this is even more so between feature-geometric and privative systems.

Therefore, different assumptions regarding the multiple parameters of feature and/or privative systems generate contrasting definitions of what a minimal diachronic change is. For instance, a given feature system may declare Bruch’s (1954) [c]-path for Cologne German velarisation non-minimal, while another arborescence harmonises with an intermediate [c]-stage, but disqualifies Martin’s (1922) glottal stop solution.

But even if there were such a thing as a consensual internal structure of speech sounds, there is reasonable doubt that actual language change is marshalled by any phonological unit at all. The quanta provided by phonological primes of whatever kind may be much too coarse in order to match reality. Movement takes place below this level, and hence can only be captured by phonetic categories. This is what Lass (1997: 222) suggests. However, he is also explicit on the fact that innovation is not infinitely finely graded either.

It is therefore useful to readjust the focus of our discussion: do diachronicians aim at writing the best record of the physical reality that once existed in the mouth of speakers? Or does their interest concern the changes that have occurred in the brain of the speakers, that is in their phonological system? As in synchronic analysis, the former activity leads to a continuum that supposes infinitely fine gradations which cannot be captured by any notation. In contrast to this phonetic point of view, phonological analysis operates with discrete units that can be distinguished and described.

Hence, it may turn out that the entire debate on the minimality of diachronic change is worthless unless the observer identifies his position: do we want to determine intermediate stages of language that were located in the mouth or in the brain? Of course, both ventures are respectable and can exist in their own right. The classical reconstructive activity has always lain on the phonetic side, and strongly so: the more phonetic detail we bring to light, the better. This is where the phonetic fantasy that was reported in section 3 comes from. Excessive phonetic atomism of this kind is also due to another factor: the only internal cause for language change (as opposed to external causes such as temperature, mountain-climbing that leads to increased breath-pressure and the like) that linguists could imagine for a long time was phonetic (analogy, which used to be called “psychological”, let aside). This is for example why Grimm’s Law was “explained” by the increase of breath-pressure. The existence of this alleged phonetic event is grounded on nothing else than the wild phonetic imagination of those who could describe, but not explain the facts.

It has now been shown that phonetic atomism is certainly unwarranted: not only is there no such thing as a minimal derivation in a continuum (a minimal solution can always be superseded by a more minimal one), but this kind of minimalism also produces absurd results (cf. section 3). Therefore, Lass (1997: 228) concludes that in many cases, the longest derivation is not likely to be the closest to what really happened.

In contrast to this phonetic point of view, the phonological stance on diachronic evolution was initiated when the difference between underlying and surface structure was formalised in synchronic analysis. The phonemic concept of the Prague School was carried into structuralist analysis and produced for example the innovative view on Grimm’s Law that was discussed earlier (Fourquet 1948; Kuryłowicz 1948). Structuralists hold that the only relevant unit is the phoneme, i.e. an object that is
located in the brain and nowhere else. Whatever its infinitely graded expression in the
mouth during phonetic translation, structurals only aim at describing the properties of
the discrete object "phoneme", as well as its relation with its peers. This is true for
synchronous and diachronic analysis alike (see for example Hockett 1965). Accordingly,
"minimal" can only mean "changing one phonemically relevant parameter at a time".
This, in turn, supposes that minimalism is defined on a language-specific basis because
phonemic systems are different in each language: a change that is phonologically
minimal in one system may not pass the minimal filter elsewhere.

Now the question arises on which side Picard's minimalism stands: is it phonetic or
phonological? One may incline to believe that Picard has a phonological understanding
of minimalism since his definition relies on feature geometry, which is a phonological
system. However, this classification does not resist examination: feature-geometric
systems are the spine of the phonological representation of speech sounds in generative
approaches, but they do not reflect phonemic oppositions at all. On the contrary, the
explicit research programme of feature geometry is to build a single system that is able
to describe all natural languages, and to capture all natural classes. Therefore, the ideal
feature geometry does not make any reference to language-specific phonemic situations
and hence is not phonological in the structuralist sense. This is not to say that
phonemic properties of individual systems could not be built into a Feature Geometry.
They probably could, and this is precisely what Clements (2001, 2002) tries to achieve
in his recent work. The point is that they are not in existing feature-geometric theories.
This actually seems to be a foundational property of generative phonological
endeavour: since its very beginnings and for reasons that may be only partly scientific,
generativists have ignored, or rather, put much energy into ignoring the structuralist
heritage: a sound is not a linguistic object in itself; it becomes one only when it enters
into a relationship with other sounds. This issue is certainly fundamental, but as far as
I can see has never been openly addressed in generative quarters. Further discussion
would take us too far afield here. In any event, there seems to be an interesting
difference in the structuralist and generative use of the word "phonological", which
(hopefully) may be overcome in the future.

In sum, thus, it appears that generative feature geometries and hence Picard face
the problem of the phonetic continuum. Grounding their venture on the identification
of natural classes of sounds, they try to define the minimal size of objects that are
phonologically relevant, i.e. the features. Unless this programme has produced a
stable result, there is no way Picard can achieve a definition of what is minimal on
generative phonological grounds. Picard's minimalism cannot rely on structuralist
phonemics either: these are language-specific, but feature geometric systems are not.
Since simple phonetic minimalism leads nowhere, it must be concluded that Picard's

principle of minimality is not operational. It does not assist the diachronic
practitioner since it is unable to tell which possible derivation deserves the minimal
crown. Moreover, it certainly does not help to fight back the arbitrary luxuriance of
asterisked forms.

4.2. There is no intrinsic rule ordering: the length of a derivation tells us nothing about
its quality
Using typological arguments in the sense of the glottalic theory supposes that
synchronous and diachronic processes are treated on a par: there are no specifically
synchronous or diachronic events. The only thing that we encounter are phonological
processes. These may have a synchronic or a diachronic manifestation and are
restricted by the possibilities and impossibilities of Universal Grammar. If there were
no phonological unity of synchrony and diachrony, there would be no point in
grounding an argument in favour or in disfavour of a diachronic process on a
synchonically observable situation.

If this principle of synchronic and diachronic unity is assumed, an interesting
parallel between the situation we are left with by the end of section three on one hand
and early generative debate on the other may be drawn. Recall the relevant questions:
how many intermediate steps do we need in order to relate an older to a modern form?
Is there any way to prefer one diachronic derivation over another on the ground of the
nature of changes? How can we avoid over-atomisation while guaranteeing the
minimal character of phonological change?

If we replace "diachronic" by "synchronic" and "an older to a modern form" by "an
underlying to a surface form", the questions asked in the preceding paragraph take us
right back to the debate that was initiated by Paul Kiparsky (1968), i.e. how abstract is
Phonology? The discussion of abstractness led to the view that abstract structures
should be dispensed with in the presence of solutions that are bound to the surface in
a more direct fashion. This position was taken because the goal of phonological analysis
is not simply to relate an underlying to a surface representation in a non-contradictory
fashion. Any analysis whose derivational path hits the attested phonetic form is not
necessarily a good analysis. The path in itself is also relevant. This point of view is
expressed in subsequent work by Kiparsky (1982b contains a number of relevant
papers) and, among others, Tranel (1981). It was the cornerstone of Natural Generative
Phonology (e.g. Vennemann 1974a, b, Hooper 1976), Natural Phonology (e.g. Stampe
1972, Dressler 1974), and also gave birth to the theory of Lexical Phonology (e.g.
Kiparsky 1982a, Rubach 1984) in the early 80's. However, the debate on abstractness
never came to a conclusive end. No discovery procedure could be established that
would allow to favour one concurrent grammar over another. Competitors differ in the
number of rules they appeal to, the phonetic distance between underlying and surface
structure as well as in the nature of the phonological processes at work. In particular,
no satisfactory definition of what is natural, simple, elegant, phonetically plausible,
psychologically real or typologically invariant could be achieved (see for example

---

16 Let me just add one thing: all that has been said in regard of feature geometry also holds for privative approaches to melodic structure (they are generative as well). Going down the same track as Clements (2001,2002), there have been (less recent) attempts at modifying the theory in a way that allows for expressing the syntactic properties of individual languages. The conceptual innovation that embodies this move is called Licensing Constraints. Relevant work has been done at SOAS in London throughout the 90s, as for example in Kaye (2001), Charotte – Göksel (1994,1996) and Cobb (1997).
be established that would be able to rank different degrees of abstractness (e.g. Kiparsky 1974, Campbell 1981, Goyvaerts 1981).17

Even though diachronicians sometimes conclude on the same indecision (e.g. Lass 1997: 228), diachronic textbooks regularly provide a catalogue of principles and conventions that ought to help the practitioner of diachronic reconstruction to find the right way to the holy asterisked form(s). Typically, naturalness, phonetic plausibility and simplicity play a prominent role here (e.g. Hock 1991: 535 ff, Lass 1997: 228 ff). The dilemma is overt, as witnessed for example by Lass (1997: 228) who provides instructions for a sound reconstruction together with doubts on their general validity on the same page.

One difference between synchronic and diachronic derivation is the fact that both forms which are related by intermediate steps may be known in the latter case (although, of course, the obligatory interpretation of texts leaves a lot of room for not knowing), whereas the underlying form and its dissimilarity with respect to the phonetic record is always subject to debate in synchrony. To all other extents, the question how arbitrary flowering of intermediate fantasy could be fought back is the same: what is abstract? What is natural? What is elegant? What is probable? What is minimal? Is a solution that calls on less intermediate steps always less abstract? Is an account which follows general typological patterns more natural?18

Quite interestingly, two opposite attitudes have emerged on the grounds of a formally identical situation. It appears that both have led to a dead-end. In synchronic analysis, the less rules a derivation required to, the "simpler", "more elegant" or "more natural" it was reputed to be (e.g. Goyvaerts 1981: 35). According to Picard's (1999) derivational minimalism, the more intermediate steps a diachronic derivation bears, the closer it matches reality. In other words, the synchronic instruction for use was to "do the job with the smallest number of rules you can", whereas Picard's diachronic motto is just the opposite.

Apart from promoting the number of rules to a criterion that allows for evaluating derivations, both approaches share the property of being "intrinsic" in the generative sense. During the period when abstractness and the evaluation measure were discussed, another question was repeatedly debated: is there a way to predict the order in which rules apply without making reference to language-specific properties? In other words, are there "families" of rules in such a way that a rule pertaining to a given family can never apply before or after a rule which belongs to another family? In spite of abundant research, no such intrinsic property of rules could be identified (e.g. Koutsoudas 1976, 1980, Kaye 1981, Goyvaerts 1981: 6). The only way of establishing the order of two rules remained the investigation of the specific system in which they occur. Making reference to information that is unrelated to the substantial properties of rules for the purpose of their ordering was called extrinsic rule ordering. Phonologists agreed that this is the only way a grammar can be discovered (the same holds true for constraint ranking in Optimality Theory).

Both cases in point here, that is "the fewer rules the better" (synchrony) vs. "the more rules the better" (diachrony), count the number of rules. Both approaches use this number in order to evaluate concurrent analyses. Since the number of the rules a derivation is made of defines a formal property that does not make any reference to the language-specific environment in which they occur, both attempts to establish an evaluation measure may be called intrinsic.

I believe that both approaches are doomed to failure for this very reason: they are intrinsic. It was mentioned earlier that no consensual evaluation measure has emerged in synchrony. More specifically, it was agreed that counting the number of rules does not determine whether a derivation is simple or complicated. In the same way, the number of intermediate steps in diachronic derivation does not reveal any simplicity or complexity per se. Moreover, in both synchronic and diachronic matters, it is not clear at all whether the simplest or most complicated derivation (in whatever sense) is the one that comes closest to reality. Rules may only be ordered extrinsically, i.e. with respect to the language-specific environment, just as much as derivations may be evaluated only on the grounds of their extrinsic properties: what do they tell us about the language, and which is the language we are talking about?

Intrinsic criteria such as Picard's minimalism are thus unsuited to refute competing diachronic analyses. In the following section, I introduce an alternative means of evaluating diachronic scenarios which does not rely on the traditional positive dialectological evidence, but on the contrary calls on the absence thereof.

5. Negative evidence: dialectal traces

Proponents of a change X > *Y > Z where *Y is induced by the principle of minimal change may seek support for their analysis in looking at dialectal evidence. A case in point has been discussed earlier in this article: lat [k] > *[ŋ] > fr [ŋ] where *[ŋ] is abundantly attested in dialects.

What would be the status of *[ŋ] in absence of any dialectal record? Does the principle of minimal change allow to assert the existence of *[ŋ] if no French, Occitan or related dialect attested affricates, and no traces thereof were found in Old French spelling and English? Wouldn't the absence of recorded affricates plead in disfavour of *[ŋ]? *[ŋ] is the supposed intermediate stage of all varieties that are subject to palatalisation of [k] before [a]. Can we believe that every single bit of the affricate reality has been independently eliminated in the hundreds of different grammars that cover the area at hand?

I submit that this kind of negative evidence is to be evaluated as a severe setback for an analysis which proposes *Y on the sole basis of its minimalism. In practice, if no trace of *Y can be found in any dialect, and if the area at stake represents a reasonable number of systems, an argument can be made against *Y.

In spite of its telling character, the absence of dialectal evidence does not necessarily lead to the elimination of the candidate analysis: the more time has gone by since a diachronic event has taken place, the more likely its hypothetical intermediate
If this scenario is correct, the following question arises: why is there not a single
trace of *[pʰ]*, *[f]* or *[pf]* in any Germanic language or dialect? Recall that Grimm's
Law is a definitional property of Germanic languages within the IE family: all and only
the languages that have carried out this change are Germanic. Just as in the case of lat
[k] > fr [], the alleged minimal intermediate stage is expected to be found in some
dialect. The fact that there is no such testimony reported from the vast territory
where Germanic languages are spoken does not plead in favour of the reality of *[pʰ]*, *[f]*
and *[pf]*.

Picard (1999) proposes the minimalist scenario(s) under (6) in reaction to my
own analysis of Grimm’s Law (Scheer 1999: 215), which interprets IE *p* > CG *f* as one
single phonological process that operates a simultaneous change of manner and
place.21 The fact that there is no trace of the alleged intermediate affricate all over the
Germanic territory does not refute its existence. However, it sits uneasily with Picard’s
minimalism. In any event, minimalism alone can certainly not refute the direct solution
IE *p(b) > CG *f*.

Now the prehistoric character of Grimm’s Law may be invoked in order to narrow
down the objection raised: too much time has gone by since these intermediate stages
were real. Traces have existed, but were evacuated everywhere over the great many
centuries. Let us therefore consider a more recent change which feeds the same
argumentation, that is the High German Consonant Shift.

5.2. The High German (2nd) Consonant Shift

5.2.1. The facts: affrication in Strong Position, spirantisation in Weak Position

Picard (1999: 71) also quotes the High German (or Second) Consonant Shift. His
presentation refers only to a subset of the facts. Let us ground the following
discussion on more complete evidence22 as well as on a review of the arguments
regarding unattested intermediate stages that have been produced for more than one
century.

Unrecorded Common Germanic voiceless stops *[p,t,k]* appear in Old High
German as affricates *[pf,ts,x]* word-initially and after Codas (i.e. also in place of CG
inated voiceless stops), while fricative reflexes *[f,s,x]* are found in post-vocalic
position (i.e. intervocally and word-finally after vowels).

This phenomenon can be inspected when comparing English to German since the
former language gives direct access to Common Germanic voiceless plosives. Table (7)
shows English forms first, followed by their modern German cognates.23

(7)  

<table>
<thead>
<tr>
<th>a. #_</th>
<th>b. Coda__</th>
<th>c. Coda</th>
<th>d. V_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>path Pfad</td>
<td>shear Schaf</td>
<td>pope Pfaffe</td>
</tr>
<tr>
<td>t</td>
<td>ten zehn</td>
<td>salt Salz</td>
<td>haste hassen</td>
</tr>
<tr>
<td>k</td>
<td>com Kern</td>
<td>streak Strich</td>
<td>make machen</td>
</tr>
</tbody>
</table>

The disjunctive contexts that govern the distribution of affricates and fricatives are called the “Strong Position” in the former case (#,C)_#, while “Weak Position” is used for the latter disjunction (V_V,V_#).24

Ségéral and Scheer (2001) call the Strong Position the Coda Mirror since its properties are exactly symmetric when compared to the Coda in both structural description and segmental effect: consonants in Codas occur before a heterosyllabic consonant or word-finally, i.e. (C,#), while consonants in the Strong Position are found after a heterosyllabic consonant or word-initially, that is (C,#)_. Consonants in Codas fall prey to lenition, while the Strong Position has the effect of preserving them from (diachronic or synchronic) damage. The High German Consonant Shift illustrates this positional effect: stops are limited to fricatives in weak positions (= Codas and V_V), but damage is limited in the Coda Mirror where their lenition takes them only to an affricate stage. At least this is the direct translation of the facts observed.

5.2.2. Interpretation of the facts: do fricatives go through an unattested affricate stage?

The crucial question is whether there has been an unattested intermediate stage where CG voiceless stops were affricates in all positions, strong and weak alike. On this analysis, a spontaneous change takes CG *[p,t,k]* to *[pf,tʃ,kʃ]* first, and then is followed by a context-sensitive move *[pf,tʃ,kʃ] > *f,s,ʃ]* that occurs only in post-vocalic position. Both processes are independent in causality and time. This scenario obeys Picard’s minimalist requirement. It is concomitant to the view that the spirantisation which produces *[ʃ,s,ʃ]* is due to the action of the preceding vowel.

Ségéral and Scheer (2001) argue for a different story: 1) affricates have never existed in weak positions, 2) no part of the High German Consonant Shift is triggered by any vocalic action, 3) both events are unitary and involve one single step, 4) the change at hand is not minimal in the sense of Picard. For the sake of clarity, both options are opposed under (8) below.25

(8)  

<table>
<thead>
<tr>
<th>CG</th>
<th>intermediate</th>
<th>Old High German</th>
</tr>
</thead>
<tbody>
<tr>
<td>b. *p</td>
<td>f / V / (C,)#</td>
<td>Ségéral – Scheer (2001)</td>
</tr>
</tbody>
</table>

On the account of (8b), 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 limited to fricatives in the latter case, but maintain an affricate status due to their positional protection in the former. If this view is correct, the unitary process at hand cannot originate in the action of a preceding vowel because even in absence of such a vocalic agent, i.e. in (C,#)_, stops are limited to affricates. As a side-effect of this analysis, the change represented by the lower branch of (8b) is not minimal in the sense of Picard because it involves a simultaneous shift in the manner and the place of articulation.

The High German Consonant Shift is of particular interest for the study of lenition and fortition since it differs from the usual pattern that is found in (diachronic) lenition processes: in Romance for instance, lenition is observed in weak, but not in strong positions. Segments that occur word-initially and after Codas appear without any damage in the modern languages.26 If no lenition occurs in the Strong Position at all, it can be argued that the process is triggered by the action of a preceding vowel. It is the fact that all CG voiceless stops do lenite (but gradually according to their position) which disqualifies the alleged vocalic trigger.

This analysis, however, crucially hinges on the nonexistence of an intermediate affricate stage in post-vocalic position. Before considering the argument coming from negative dialectal evidence, let us see how the issue of intermediate asterisked stages has been addressed in the literature on the High German Shift for over a hundred years.

---

23 General presentations of the High German Consonant Shift may be found in, among others, Braunig – Eggers (1987: 81ff), Lerchener (1971: 13ff, 136ff), Meisen (1961: 76ff), Lessjak (1913: 152ff), Proksch (1939: 78ff), Wilmanns (1911: 54ff), Bentz (1969: 65ff, 1971: 147ff, 1975: 79ff), Zauche (1974: 131ff, 1987: 94ff), Sonderegger (1974: 155ff). The behaviour of Common Germanic stops in internal Codas (=before a heterosyllabic consonant) may not be controlled because all relevant configurations were affected by the First Consonant Shift (Grimm’s Law), thus lat capus, noces, rectus = OHG haft, naut, raut. Note that the fricative coming from CG *[k]* is invariably represented by “ch” in spelling, but has become subject to contextual influence subsequently to the process at stake: “ch” = *k* after *o, u, e*, while [k] occurs after front vowels. Spell “x” = *[k]*. The affrication of CG *[k]* to *[kʃ]* has occurred (or survived) only in High-Alemmanic (Switzerland); the simple stop is found (or has been restored) elsewhere, thus [korn] and [dakon] in NHG.

24 These terms are chosen according to the bearing they have on consonants. Based on independent but phonologically identical phenomena, both the German and the Romance tradition have come to this conclusion and terminology. See for instance Lerchener (1971: 143) “phonetisch schwache vs. phonetisch starke Stellung”, Pope (1952: 96), Bourciez (1926: 147).

25 As for Grimm’s Law, the labial series alone represents the Tenseverschiebung in the following discussion. Everything that is said for labials also applies to dentals and velars.

26 For example, Latin obstruents are faithfully restored in French if they occur in the Coda Mirror, whereas they undergo various lenition processes in weak positions (compare e.g. lat porta, tajsa vs. rupata, lupata, ripa vs. fr porte, taupe vs. route, boup [bup], rive). The same holds true for the evolution of Iberian–Romance sonorants. This kind of classical data is discussed at length in Ségéral – Scheer (2001). Bourciez (1926), Pope (1934) for example expose the French situation.
5.2.3. Arguments for and against an intermediate aspirated stage

The discussion of the precise relation between CG *p and the recorded OHG $f/$, f has been largely influenced by two properties whose linguistic character may be doubted. For one thing, it was taken for granted that Grimm’s Law and the High German Consonant Shift are manifestations of the same fundamental force. It is certainly true that the parallels are striking (spirantisation, devoicing of voiced consonants, unshifted *p after obstruents in both cases), but there are also numerous contrasts: 1) *t in *tr moves with Grimm, but remains stable in High German, 2) spirantisation of *t produces [?] with Grimm, but [?] in High German, 3) there is no dialectal variation at all among Germanic in the case of Grimm, whereas the High German area is scattered with isoglosses that originate in the Second Shift, 4) and most prominently, Grimm is a spontaneous sound shift, while the Second Mutation obeys a contextual conditioning: affricates in strong, fricatives in weak position.27 Be that as it may, the unity of both mutations has scarcely been questioned (though examples are Paul 1874: 191, Heinertz 1925: 65, Güntert 1927: 4, Brinkmann 1941, FOURQUET 1948: 50, 91). The general opinion holds that “both mutations are phases of a continuous sound change that originates in a uniform phonetic tendency” (Prokosch 1917: 3).28

For these reasons, the analysis of the High German Shift was built on the model of Grimm’s Law. This is specifically true for an alleged aspirated intermediate stage: since it is generally accepted that IE *p underwent (spontaneous) aspiration before Grimm applied (and that aspiration was the actual cause of spirantisation, cf. section 2.2), there must have been an aspirated intermediate stage between CG *p and OHG $f$ as well. This line of thought is more or less explicit in Braune (1874: 45ff), Meyer-Benfey (1901: 109), Wilmanns (1911: 55), Lessiak (1933: 13, 152ff), Nordmeyer (1936), FOURQUET (1948: 92), SCHIRMUSKI (1962: 350), PENZL (1971: 151), SONDeregger (1974: 163f, 1979: 131f).

The second factor that is not really linguistic but has dominated the debate is simple prestige and authority: the founding article on the High German Shift was written by Wilhelm Braune (1874) and published in the first issue of the central neogrammarian organ Pauls and Braune’s Beitrag. On the grounds of a pioneering exploration of the very scant dialectal material that was available in these early days, Braune (1874: 45ff) concluded that CG *p went through both an aspirated and an affricate stage on its way to OHG $f$ in post-vocalic position: CG *p > *ph > *pf > OHG $f$. Ever since, this scenario is reproduced in textbooks and articles as the “classical” picture without questioning or arguing (e.g. Wilmanns 1911: 55, Penzl 1969: 65f, Szule 1974: 134, Metzke 1993: 109).

Apart from the analogy between both mutations and Braune’s authority, which are the actual arguments for 1) aspirated and 2) affricate intermediate stages? Let us first address the former issue. The kind of “phonetic” reasoning that prevailed until way after the Second World War was largely intuitive and impressionistic (see Grimm’s Law above). Since a relation had to be established between a [p] and an [f], there could be no doubt that the process at hand was a weakening. Hence, the candidates for intermediate steps were articulations that are slightly weaker than [p] but yet stronger than [f]. Even though aspirates are often considered stronger than plain voiceless stops, [ph] was viewed as a satisfactory solution. Intuitively, people thought that it was “physiologically” impossible to produce a direct transition [p] > [f] without intermediate [ph] (e.g. Vonficht 1964: 27 for a critical review of this “phoneticism”). Hence, the alleged aspirated stage, on this “phonetic” line of thought, is little more than a result by default which is grounded on the belief that a direct transition is “physiologically” impossible.

A more serious argument in favour of an intermediate aspirated stage is the presence of aspirated voiceless stops in many (though not all) modern dialects. The modern aspirated testimony is then interpreted as an unaltered continuation of the alleged intermediate step. The argument crucially hinges on the distribution of affricates in the area that has been touched by the High German Shift (i.e. South of the Brennath-line). As may be seen on the map under (9), the present-day dialectal division is entirely based on the increasing absence of affricates and fricatives towards the North.

---

27 Lercher (1971: 19) provides more detailed discussion of the (dis)similarity of Grimm and the High German Shift.

28 Along these lines for example Scherer (1878: 121f), Lessiak (1933: 152ff), Schmitt (1949: 17f), Foley (1970: 12), SONDeregger (1979: 136f), Heinertz (1925), Kasten (1939, 128ff) and Lercher (1971: 181, 32f) offer a general overview and more literature. The analogy between Grimm’s Law and the High German Shift is also used to have a non-linguistic motivation. Following Jacob Grimm, some authors believed that the resident property of the Germanic character is to show wild, juvenile will and innovative courage. This character surfaces from time to time (especially in periods of migration) and then provokes profound changes in the language (e.g. Mistka 1951, 1952; Hölder 1957, 1958).
Only High Alemannic possesses the full set of affricates (labial, dental, velar) in all relevant positions (i.e. word-initially, after sonorant Codas and in place of former geminates). The velar affricate is already missing in other varieties of Oberdeutsch, i.e. Bavarian and Swabian. When moving further to the North, the distribution of affricates is more and more restricted according to both parameters, i.e. the place of articulation ([kʃ] being absent anyway, [f] is present everywhere, while [pʃ] is rarer and rarer) and the position in the string (word-initial stops resist better against affrication than their geminated and post-Coda peers). This produces modern dialectal systems where affricates cohabit with aspirated stops, e.g. [pʃ, tʃ, kʰ] in Middle and Northern Bavaria as well as in Swabian, [pʰ, tʃ, kʰ] in Rhein- and Moselfränkisch. The general decay of affricates towards the North laid aside, the distribution of affricates and aspirated stops in the particular dialects is random and unpredictable.

The argument for an intermediate aspirated stage is thus obvious: the modern aspirated stops failed to undergo affrication and therefore represent an older layer (e.g. Lessiak 1933: 13, 153). However, this reasoning is objectionable on two grounds. For one thing, the development of aspiration can be secondary as well, i.e. younger than the High German Shift. There is no necessary concomitance of the High German Shift and aspiration since aspiration also occurs on lands that were not affected by the Second Mutation. For example, English is a language that has developed spontaneous aspiration of voiceless stops, and nobody will want to correlate this fact with the High German Shift.

Yet another solution for the presence of aspirated stops in the modern dialects is proposed by Vonficht (1958, 1964), who draws attention to the presence of aspirated stops North of Benrath and also to the absence of these articulations in quite a few systems that did undergo the High German shift (e.g. Ripuarian). According to him, there has never been an intermediate aspirated stage. The existence of modern aspirated stops South of Benrath is due to the action of two movements that are opposite in kind and geographic dissemination: aspiration comes down from the North, while affrication moves northwards. In the midlands, both waves contaminate each other, resulting in systems with random distribution of aspirated and affricate articulations. Vonficht (1958, 1964) points out that an aspirated intermediate stage supposes the abandonment of aspiration in non-aspirated areas South of Benrath once the High German Shift was completed. This, he argues, is not any more plausible than the secondary development of aspiration due to a spontaneous movement or to Northern contamination.

Finally, an argument in favour of aspiration is made on the grounds of the so-called exceptions to the High German Shift, i.e. CG voiceless stops that have resisted the mutation. Voiceless stops were protected against any change in case they followed an unobstruent. Remarkably enough, this is exactly the context in which Grimm's Law was inhibited (see the examples in note 23). CG Voiceless stops could appear after the obstruents [s, f, x] as for example in CG *stainaz, kraftiz, naxtiz > OHG stein, kraft, nacht "stone, strength, night", and indeed were carried into OHG without any modification. The argument that builds on this fact actually combines the three reasons that have been mentioned thus far: analogy with Grimm's Law, phoneticism and typological plausibility. Since the parallel with Grimm's Law is so striking (and this time really linguistic), it was taken for granted that the non-mutation in both shifts is due to the same cause. And the one that was traditionally advocated for Grimm's Law is of "phonetic" kind (e.g. Penzl 1969: 45, 1975: 52): voiceless stops do not develop spontaneous aspiration because the reinforced breath-pressure that causes aspiration in other contexts is "diluted" or "deviated". That is, a non negligible portion of the brutal airflow is absorbed by the articulation of the preceding obstruent, which is also an active breath-consumer. Hence, post-obstruent stops remain unaspirated and therefore are not subjected to Grimm's Law or the High German Shift since these processes only operate over aspirated stops. Authors that argue for an intermediate aspirated High German stage along these lines include Wilmanns (1911: 722), Szülc (1987: 98), Penzl (1971: 151, 1975: 81), Paul et. al (1989: 116), Mettke (1993: 109).

Finally, Penzl (1971: 151) points out that the spontaneous aspiration of voiceless stops that can be inspected in Modern English also leaves out those stops that occur after obstruents, e.g. sport, not *spört.

On the other hand, a number of voices argue against this so-called Behauensgstheneorie. The point made by Vonficht (1958, 1964) has already been reported. Hammerich (1955: 155ff) also denies an intermediate aspirated stage. In his opinion, the triggering factor of the High German Shift was the gemination of CG simplex voiceless stops, which then moved directly to affricates on one hand and spirants on the other. Authors that argue against aspiration include Scherer (1968), Schatz (1927: 95), Brinkmann (1941: 85) and Schützeich (1961: 238ff).

Whatever the correct solution, it has only indirect bearing on the question that is of interest here. Recall that the two concurrent scenarios under ($) differ only in regard of the events in post-vocalic position: either affricates did appear in this location (Picard 1999), or they did not (Ségéal – Scher 2001). In case they did, there was a so-called unitary affricate stage (affriziertes Einheitstadium), which means that all CG voiceless stops had shifted to affricates in all positions. This solution is minimal in the sense of Picard. On the other hand, if affricates have never existed after vowels, CG voiceless stops have become fricatives directly, and the change is not minimal (place and manner are altered). As a matter of fact, both scenarios can live with an intermediate aspirated stage that preceded actual affrication and spirantisation. Whether CG voiceless stops were aspirated prior to the other changes or not does not change any parameter of the discussion. The only crucial question concerns the existence of affricates in post-vocalic position. Nevertheless, it is worth knowing about the debate on aspirated stops since most of the defenders of a unitary affricate stage suppose a necessary aspirated basis.

---

29 This is attested both by modern dialects and OHG script: Ostrich, who represents Rheinfränkisch, writes affricates after sonorants and in gemination, but preserves unaspirated plain stops word-initially. On the other hand, the Ostfränkisch Talian uses affricate graphemes in all positions (e.g. Penzl 1969: 69).
5.2.4. Arguments for and against an affricate intermediate stage

 Braun (1874: 47ff) himself discusses another early view on the matter: Scherer (1868, 1870: 265) holds that CG voiceless stops were directly spirantised in post-vocalic position without any intermediate affricate stage. Braun explicitly accepts this option as a possible scenario, but prefers his own solution: CG *p > *pf > *pf > OHG f. He favours an intermediate affricate stage because of two reasons: the attested change OHG pf > f in certain words, and the gemination of the resulting fricatives (Engl hate = germ haben). As far as I can see, no other substantial arguments have been adduced in order to establish the alleged unitary affricate stage in more than 120 years of debate. Since Braun's pioneering work, most of the grammars just reproduce his view, which has therefore become authoritative (e.g. Wilmanns 1911: 55, Penzl 1969: 65f, Szulc 1974: 134, Mettke 1993: 109). Braun's arguments are sometimes exposed, but no new evidence has joined in. Let us therefore review Braun's reasoning in some greater detail.

The evolution of OHG manuscripts witnesses the simplification of affricates in post-Coda position. Braun (1874: 48) mentions words such as got waipan = OHG wē rávan = wē rāvan, got. hilpan = OHG hēlpän; hēlfän (Penzl 1975: 82 for example offers more material). OHG pf after liquids is regularly attested in manuscripts older than the 9th century, but is further shifted to OHG f from that period on (see Franck 1909: 103f for Franconian scriptural evidence). While Braun and many of his followers (e.g. Eggers 1987: 120, Szulc 1974: 95f) just say that (labial) affricates become fricatives in "some" words or after [r,l], Wilmanns (1911: 67f) has discovered the regularity that controls the change: deaffrication occurs everywhere except when the preceding sonorant is homorganic. Possible sonorants in this position are [m, n, l]. The dental OHG t is remains entirely untouched because it always shares the place of articulation with the preceding sonorant (there are no words with -nto-). The labial OHG pf is maintained after the labial [m] (lat. campus > OHG kampf > NHG Kampf "fight"; OHG scimpf > NHG Schimpf "shame") but becomes [f] when preceded by the dentals [r,l] (there are no words with -npf-): [r,l] OHG dorf > dorf > NHG Dorf "village"; [l] OHG hēlpän > hēlfän > NHG helfen "help" (Wilmanns 1911: 68f provides rich material). The situation of the velar affricate is difficult to control since its existence is more or less restricted to Alemannic. However, modern dialects seem to attest the expected distribution, i.e. [kʃ] after [s], against [ʃ] after [l].

The conservative option of homorganicity on consonant clusters is rather convincing since it may also explain the aforementioned absence of affrication in CG *tr clusters (Wilmanns 1911: 74, e.g. got trudan = OHG trētan > NHG treten "to step"); all other sequences of rising sonority CG *pl, pr, kl, kr (there is no *nl) do not share the place of articulation and are therefore prone to lenition (see Honeybone 2001, 2002 for a more general interpretation that goes beyond Germanic).

The argument that is produced on the grounds of this evidence, then, appeals to the precedence of the evolution affricate > fricative: if deaffrication has occurred after Codas, we know that this is something the language allows for. Hence, we may as well assume that there was an unrecorded affricate stage after vowels, and that these post-affricates were further transformed into fricatives.

This type of argument has been mentioned earlier: a putative asterisked intermediate stage is plausibilised by the existence of an identical or similar process in the same, a neighbouring or an unrelated language. Clearly, this is no positive evidence that hints at the existence of affricates in post-vocalic position. Rather, it qualifies this option as one possible scenario.

Braun's (1874) second argument is the mysterious gemination that accompanies post-vocalic spirantisation: Engl open = OHG offen > NHG offen; water = OHG wasar > NHG Wasser; Engl break = OHG brēðan > NHG brechen (OHG spelt geminates are pronounced only simplex). The real geminate value of the doubled consonants is beyond any doubt. For one thing, a subsequent process reduced the geminates to simple fricatives after long vowels and word-finally. Hence, forms such as OHG slāfän, slaffan > NHG schlaffen "to sleep" (<sch> represents a long vowel) cohabit in manuscripts, whereby the geminated option was progressively abandoned. The word-final degemination is apparent in alternations produced by inflection such as OHG NOMsg skif (= got skip) vs. GENs skiffs (= got skipis), e.g. Lessiak (1933:153f, Prokosch (1939: 81). Degemination in these positions is a perfectly reasonable process, and the simple fact that manuscripts consistently distinguish between simplex and double writing supposes that both symbols correspond to two different sounds. Also, MHG verse shows that the surviving spelt geminates had actual geminate value (Lessiak 1933: 153f, Braun 1874: 48, Michel 1979: 109f, Mettke 1993: 116). Finally, many modern dialects such as Silesian, Middle- and Northern Bavarian have preserved geminates until today.

Hence, post-vocalic spirantisation is not suspect in any way, but there is no apparent reason why simplex stops should produce geminate results. If it is assumed that the fricatives at hand were native in affricates rather than in stops, Braun argues, the gemination may be explained: affricates are twofold. They may give birth to geminate fricatives when their fricative part assimilates the preceding stop; pf > f. In other words, the intermediate affricate guarantees the transition from a simplex to a twofold articulation.31

It is quite obvious that the equation of graphically and phonetically twofold affricates with true geminates is erroneous. Everything we know about affricates in modern German and other languages shows that they pattern with simplex consonants, not with geminates. It would be quite bewildering indeed if OHG affricates had geminate value, i.e. occupied two skeletal slots. We would then have to admit the

30 A few isolated affricates have survived after liquids into MHG and even NHG, see for example Wilmanns (1911: 68f), Szulc (1987: 96). Recurrently quoted words are lat. carp > OHG karp > NHG Karpfen "carp": lat. hāra > OHG hāpfa, hārfa > NHG hāpfe, hārfe > NHG Harfe, Luther still writes <Harpe> = "harp": MHG kēlpf "high-spirited", extinct in NHG, MHG kēlpf > NHG scharf.

31 In addition, many authors such as Lessiak (1933: 153f) believe that in turn, an affricate can be related to a stop only via an aspirated stage for articulatory and other phonetic reasons.
existence of word-initial geminates, and even worse, of word-initial geminates followed by another consonant: OHG pfegan, pfluog, pfluonta > NHG plegen, Pflug, Prünke “to look after, plough, prebend”. Fourquet (1948: 80f) points out another very obvious fact that refutes a twofold value for affricates: if affricates were already geminate, they could not possibly be further geminated. However, geminated affricates precisely are the result of CG geminated stops: old saxon appul = OHG appful > NHG Apfel “apple” (see for example Wilmanns 1911: 65, Meisen 1961: 77 for the interpretation of the graphic evidence).

The gemination of spirantisated post-vocalic CG stops thus remains mysterious. In no event can they be derived from the twofold identity of affricates. Hence, the gemination of fricatives cannot ground an argument in favour of the existence of unrecorded affricates in post-vocalic position.

There are also voices that explicitly reject the alleged unitary affricate stage. Fourquet (1948: 80f, 91f) has already been mentioned. Schmitt (1949: 20f) casts doubt on the existence of post-vocalic fricatives from the phonetic point of view. Sacht (1927: 95) also argues for a direct shift from CG voiceless stops to fricatives without intermediate affricate stage.

Hammerich (1955) proposes another account for the geminates result of spirantisation in post-vocalic position: he argues that the CG voiceless stops *p,t,k have first undergone gemination in post-vocalic position. Hence, the High German Shift did not operate on a uniform series *p,t,k, but on simplex stops in strong vs. geminates in post-vocalic position. The different result, i.e. affricates in the former, against geminated fricatives in the latter context, simply continues the original system. On this account, of course, affrication and spirantisation are two distinct processes that operate on different inputs. A unitary affricate stage including post-vocalic instances is out of the question.

Lerchner (1971: 144) also concludes that the geminated character of post-vocalic fricatives does not call for a unitary affricate stage: “Für die Doppelspiranten ist es nach der dargelegten Einsicht in den phonetischen Mechanismus des Wandelns nicht erforderlich, eine affrikatische Zwischenstufe anzunehmen. Beide Lauttypen sind voneinander unabhängige [...] Entwicklungen aus aspirierten Allophonien unterschiedlicher Verschlußfestigkeit. Diese unterschiedliche Verschlußfestigkeit ergibt sich [...] aus den phonetischen Konsequenzen verschiedener Positionen, ‘schwache’ gegen ‘starke’ Stellung.” This is exactly the position embodied under (8b) (Ségéral – Scheer 2001): the split of voiceless (aspirated) stops into affricates and fricatives is a direct consequence of the position in which they occur. We face one single phonological process that produces two different results according to a given context, but no intermediate (affricate) stage.

Finally, an argument against the existence of affricates in post-vocalic position can be made on the grounds of the so-called Ripuarian residual words (“Restwörter”), e.g. Braune – Eggers (1987: 155). As anywhere else South of Benrath, CG post-vocalic voiceless stops appear as fricatives in this most Northern mutation-area (e.g. Cologne): CoIG esse [esə] = engl eat “eat infinitive”, CoIG us = engl out. However, there are a few words that bear the unshifted dental, both in OHG manuscripts and modern dialects: OHG Rip dat, it, wat, allet = NHG das, es, was, alles “that, it, what, all”. This phenomenon is not restricted to the word-final position (as Paul 1879: 554ff thought), nor does it concern only “little” or “grammatical” words: it recurrently opposes the (shifted) infinitive and (unshifted) inflected participial forms of weak verbs (e.g. Meisen 1961: 78): OHG Rip inf groessen, bosen (= engl greet, old engl boocan = NHG grüßen, büsen “to greet, to do penance”) vs. 1st sg groote, boote, past participle gegroet, geboot. These residual words remain entirely mysterious today. Of course, the easiest explanation for their existence is to interpret them as intrusions from the North. If on the other hand they are viewed as true unshifted residues of CG voiceless stops, we hold in hand the origin of spirantisation: a plain voiceless stop, not an affricate.

The philological debate reported shows the diversity of views. Since this phenomenon has been studied by Braune in 1874, the existence of post-vocalic affricates is based on two and only two arguments, one of which is clearly wrong (affricates are not and have never been geminates). The other shows that an evolution affricate > fricative is 1) possible in general and 2) exists in the same language under certain contextual conditions. It does not constitute any positive evidence. On the other hand, the promoters of a direct shift from (aspirated) stops to fricatives mostly try to show 1) that the two arguments quoted are unfounded and 2) that there is no necessity for post-vocalic affricates on any other grounds (phonetic, systemic). However, it is also true that nobody has adduced evidence that refutes the possibility of a unitary affricate stage.

Hence, can the competition between (8a) and (8b) be decided on other grounds? I submit in the following section that negative dialectal evidence may lend an argument into the discussion that has not been previously considered, and which puts short any phonetic, graphic and analogical debate.

5.2.5. There is not a shred of evidence for post-vocalic affricates in any old or modern dialect

It is not reasonable to discard the solution that does not recognize an intermediate affricate stage in weak positions for the simple sake of being non-minimal. I have tried to show in the previous section that minimalism, just as naturalness, elegance, simplicity and the like, is not a neutral, objective and measurable notion: it depends on the phonetic detail that each linguist is willing to represent (Fouche) and the particular system of melodic representation adopted (Picard). Moreover, minimalism does not guarantee a correct analysis: two concurrent diachronic derivations may be each perfectly minimal, and still at least one of them must be wrong.

It seems to me that a better instrument is provided by dialectal evidence. In this field, both scenarios (8a) and (8b) make contrasting predictions that may be confronted with fact: if affricates were general and really occurred in weak positions, we expect to find traces of this affrication in word-final and intervocalic locations somewhere in the huge and extremely well-studied High German dialectal space. If on the other hand CG voiceless stops directly moved to fricatives in weak positions, affricates have never existed intervocally and word-finally. Accordingly, no traces thereof are expected to be found in the relevant dialects.

The dialectal testimony is entirely unambiguous: over the complete area at stake, no affricate trace in weak position is attested. This negative result is certainly not due to
The evidence presented in this section is of negative nature ("there is no X" instead of "there is X"). This is unusual in dialectology and diachronies, but not in synchronic analysis. I submit that it casts serious doubt on the diachronic derivation that goes through a generalised affricate stage. This negative evidence is even more striking when the dialectal testimony of the High German Consonant Shift is compared to the modern traces of other processes which have occurred at about the same time. One case in point is the aforementioned evolution of lat [k] > *[f] > a (> French [ʃ]). This assimilation is usually dated back to the 5th century (e.g. de la Chaussée 1989: 67), and it has left abundant dialectal traces. Is it reasonable to assume complete extermination of the affricate stage in one case, against its massive dialectal presence in the other?

6. Conclusion

Everybody agrees that two attested evolutionary stages which involve a large phonetic distance are related by one or several intermediate forms. The purpose of this article was to inquire on the precise definition of "large" in this context. I have attempted to show that a formulation in terms of pure phonetic criteria may not be achieved. Phonetic determination easily (or rather, inevitably) leads to over-atomisation of intermediate stages. It is quite safe to say that these do not describe any reality other than the fantasy of their author (section 3). Also, it is an illusion to believe that there is something like a minimalistic crown in the phonetic continuum.

Nobody knows what "minimal" exactly is. We are able to tell that a given change is not minimal enough (cf. lat [k] > French [ʃ]), and that another is excessively minimal (cf. Fouc'hé). But where exactly do we draw the red line? Unfortunately, the only thing on which we can ground our judgement is experience and intuition. Recurring to allegedly objective criteria such as feature geometry (Picard 1999) is of no salvation because there is no such thing as the distinctive feature system: controversial proposals coexist in ongoing debate. Moreover, there is reason to believe that a descriptive unit such as a distinctive feature is too coarse and does not match the actual phonetic sound change that occurs in the mouth. The only known way that allows us to elude the phonetic continuum is to recur to discrete phonological units, i.e. phonemes. This, however, obliterate any hope for a universal criterion since phonemic systems are different across languages. In sum, Picard's minimalism may not be grounded on discrete phonological units of either structuralist or generative kind, and phonetic solutions do not qualify because they inevitably lead to atomistic fantasy. Therefore, the desirable universal criterion of minimalism is not operational.

Diachronics are exposed to the same kind of uncertainty when several concurrent diachronic derivations run against each other. Cases may arise where no phonetic or phonological minimalism helps deciding in favour or in disfavour of either competitor. Indeed, there are multiple ways of being minimal, and hence incompatible derivations may relate a given input to a given output through minimal intermediate steps. A good deal of the generative phonological discussion in the 70s and early 80s has stemmed from the very same dilemma in synchronic derivation: how do we decide between different competitors when surface forms are derived from underlying
representations? In presence of incompatible grammars that possess the same
descriptive and empirical adequacy, no uncontroversial evaluation measure could be
brought to light. Interestingly, nothing that can be assimilated to the derivational
minimalism, which is fashionable in diachronics, has been proposed. Rather, arguments
invoking elegance or efficiency that lead to “short” derivations have been advocated. I
am not aware of a voice saying that the best grammar is the one which contains the
greatest number of rules. Hence, it seems that synchronic and diachronic analysis have
reached opposite conclusions when faced with a formally identical problem, that is the
evaluation of concurrent derivations. Another parallel is the fact that criteria such as
simplicity or elegance have not led to an acceptable evaluation measure in synchronic
analysis either. I have suggested in section four that derivational minimalism in the
sense of Picard (1999) must fail in diachronic analysis because it shares a formal
property with simplicity etc.: both criteria are of intrinsic nature. They disregard the
specific content of each proposal, as well as the language-specific environment in
which they operate.

Independently of this issue, it is to be noted that diachronic derivation has an
invaluable advantage over synchronic activity: it may recur to dialectal evidence.
There is no such thing as a range of contrasting surface forms that stem from the
same underlying representation. I have tried to demonstrate how both positive and
negative dialectal evidence may qualify or disqualify a particular diachronic scenario.
If the former kind is deeply rooted in any reconstructive activity since forever,
grounding an argument on the absence of a given form in the relevant dialectal area is
less common. I submit that negative dialectal evidence may enlighten diachronic
events as well.

Finally, I believe that the two competing views which are debated in this paper
represent two fundamentally different approaches to sound: requiring diachronic
change to be minimal is phonetic in essence, whereas allowing for non-minimal shifts is
phonological. The former holds that “processes represent responses to phonetic
difficulties [...] , and each process makes substitutions by altering a single phonetic
property to remedy the difficulty” (Donegan – Stampe 1979: 136f., quoted by Picard
1999: 68). This attitude shines through Picard’s entire text, traces thereof being for
example the repeated reference to phonetic statistics (“[A] occurs X times more often
in natural language than [B]”) and the phonetic definition of what a natural sound
change is (Picard 1999: 68, with reference to Ohala 1974).

The phonological viewpoint does not deny the existence of phonetic motivation in
sound change. Only does it credit a subset of the diachronic processes observed to
things that happen in the brain, rather than in the mouth. This position appears to be
trivial: it represents the foundations of 20th century linguistics. Since Saussure, there is

34 Sociologically differentiated expression could be said to instantiate this kind of situation. However,
there is no reason a priori to consider that different pronunciations in different social contexts amount to
one single underlying form. But even if this were the case, the linguist has no direct access to a set of
surface-forms that belong to a given sociologically conditioned expression. By contrast, dialects provide
direct testimony of different diachronic outputs which amount to one single primitive form.

a level of language that exclusively obeys grammar-internal mechanisms (= Language); this
level is different from the one where external influence such as phonetics, sociology etc. come in (= Parole). The existence of an autonomous and self-contained
level of grammar has later been taken over by the generative approach (I-language, competence vs. E-language, performance). Anderson (1981) for example fleshes out the
Saussurian stance in phonology.

The Saussurian point of view, as far as phonology is concerned, has recently been
challenged from the phonetic quarters (e.g. Coleman 2002) and by neo-behaviourists
(e.g. Bybee 2001, Carr 2000). These authors deny that there is any phonological
decision made in the brain without recourse to extra-grammatical parameters. This
attitude comes quite close to the (neo)grammarians) phonetic conception of diachronics,
of which derivational minimalism is an outgrowth.

The debate regarding minimalism may thus probably be reduced to the
philosophical opposition between empiricism and rationalism (mentalism): if one
believes that language and grammar are no more than the sum of non-grammatical
constraints on human communication, phonetic minimalism à la Picard makes sense or,
actually, follows. If on the other hand one stands on Saussurian and Chomskian
grounds, holding that there is something like an autonomous phonology, the manner
and the place of a sound may well be altered in the course of a single phonological
event, provided that the origin of this event is not located in the mouth, but in the brain.

Following Saussure, then, there seems to be little reason to adopt contrasting
attitudes towards synchronic and diachronic derivation. Both may originate in
properties of the mouth and properties of the brain. In both cases, the red line between
the two kinds of motivation is sometimes hard to draw. And there does not seem to be
a way to evaluate competing derivations on purely formal and objective, that is intrinsic
grounds: criteria such as minimality, elegance, naturality, simplicity, psychological,
typological or phonetic plausibility are not measurable in all cases, and their strict
application may lead to outlandish results. Therefore, (internal) reconstruction should
strive towards “minimal” sound change in absence of qualified counter-evidence. However, non-“minimal” solutions must be accepted without hesitation if good reasons
plead in their favour.

The common practice of diachronic reconstruction partly recurs to experience and
intuition. Some may feel that this is unsatisfactory in the light of the scientific ambition
of linguistics. However, there is little chance of developing a method that is grounded
on objective and measurable criteria alone. Rather, there is good reason to believe that
this state of affairs is the common, natural, useful and necessary condition of science.
Chomsky (1965: 20) asks “whether the important feature of the successful sciences has
been their search for insight or their concern for objectivity. The social and behavioral
sciences provide ample evidence that objectivity can be pursued with little consequent
gain in insight and understanding. On the other hand, a good case can be made for
the view that the natural sciences have, by and large, sought objectivity primarily insofar
as it is a tool for gaining insight.” In other words, minimality is a desirable result, but a
bad counsellor: “reason and science go often different ways. A cheerful anarchism is
also more human. It is better equipped for achieving progress than conceptions of ‘law and order’” (Feyerabend 1986: 13).

TOBIAS SCHEER
2274, chemin de la Treille
(chemin de Peylouber)
06740 CHÂTEAUNEUF DE GRASSE
FRANCE

REFERENCES

Anderson, John – Charles Jones

Anderson, John – Colin Ewen

Anderson, Stephen

Anglade, Joseph

Aubin, Hermann – Theodor Frings – Josef Müller

Bach, Adolf

Bertram, Otto

Bourciez, Edouard

Braune, Wilhelm
1884 “Gotisch DDJ und altmordisch GGJ”, Beiträge zur Geschichte der deutschen Sprache und Literatur 9: 545-548.

Braune, Wilhelm – Ernst A. Ebbinghaus

Braune, Wilhelm – Hans Eggers

Bretschneider, Anneliese

Brinkmann, Henning

Bruch, Robert

Bruck Anthony – Robert Fox – William La Caly (eds.)
1974 Papers from the parasession on Natural Phonology. Chicago: Chicago Linguistic Society.

Burdach, Konrad (ed.)

Bybee, Joan

Campbell, Lyle

Carr, Philip

Charette, Monik – Asli Gökşel

Chomsky, Noam

Clements, George – Elizabeth Hume

Clements, George

Cobb, Margaret
Coleman, John  

Collinge, N. E.  

Coetsee, Frans Van – Herbert L. Kufer (eds.)  

Cyran, Eugeniusz (ed.)  

Dalbera, Jean-Philippe  

Dell, François  

Dingwall, William Orr (ed.)  

Dinnsen, Daniel  

Dinnsen, Daniel (ed.)  

Donegan, Patricia – David Stampe  

Dressler, Wolfgang  

Dziubalska-Kolczynska, Katarzyna (ed.)  

Feyerabend, Paul  
1986 *Wider den Methodenzwang*. Frankfurt am Main: Suhrkamp.

Foley, James  


Fouche, Pierre  

Fourquet, Jean  


Franck, J.  


Frings, Theodor – L. E. Schmitt  

Frings, Theodor  

Fujimura, O. (ed.)  
1973 *Three dimensions of linguistic theory*. Tokyo: TEC.

Gamkrelidze, Thomas – Vjačeslav Ivanov  

Goldsmith, John (ed.)  

Goyvaerts, Didier  

Goyvaerts, Didier (ed.)  

Gribble, Charles E. (ed.)  
1968 *Studies presented to professor Roman Jakobson by his students*. Cambridge, MA: Slavica.

Güntert, Hermann  

Hall, Tracy (ed.)  

Hammerich, L.L.  
1955 “Die germanische und die hochdeutsche Lautverschiebung”, *Beiträge zur Geschichte der deutschen Sprache und Literatur* 77: 1-29, 165-203.

Heike, Georg  

Heinertz, Otto  
Heinrichs, Heinrich Matthias  

Hellberg, Staffan  

Hirt, Hermann  

Hock, Hans Heinrich  

Hockett, Charles  
1965 “Sound change”, Language 41: 185-204.

Höfler, Otto  

Honeybone, Patrick  

Hönig, Fritz  

Hooper, Joan  

Hulst, Harry Van Der – Norval Smith (eds.)  
1982 The structure of phonological representations I. Dordrecht: Foris.

Iversen, Gregory – Joseph Salmons  

Jung, Edmond  

Karras, M. C.  

Kardela, Henryk – Bogdan SzymaÅ†ek (eds.)  

Kaye, Jonathan  


Kaye, Jonathan – Jean Lowenstamm – Jean-Roger Vergnaud  

Keller, R.E.  
1961 German dialects. Manchester: Manchester University Press.

Kenny, Michael – Charles Kisseberth  

Kiparsky, Paul  

1982 “From Cyclic Phonology to Lexical Phonology”, in: Harry Van Der Hulst – Norval Smith (eds.), 131-175.


Kluge, Friederich  

Knoop, Ulrich et al. (eds.)  

Koutsoudas, Andreas  

Koutsoudas, Andreas (ed.)  

Krahe, Hans – Wolfgang Meid  

Kuepper, Karl J.  
Kuryłowicz, Jerzy  

Lakoff, George  

Lass, Roger  

Lerchner, Gotthard  

Lessiak, Primus  

Lightner, Theodore  

Martin, Bernhard  

Martin, Roland  

Martinet, André  

Meillet, Antoine  

Meisen, Karl  

Wolfgang Moelleken (ed.)  

Metzke, Heinz  

Meyer-Benley, H.  

Michels, Victor  

Mitzka, Walther  
1943 *Deutsche Mundarten*. Heidelberg: Winter.

Moulton, William  

Müller, Josef  

Münch, Ferdinand  

Nordmeyer, George  

Ohala, John  

Paul, Hermann  
1874 “Zur Lautverschiebung”, *Beiträge zur Geschichte der deutschen Sprache und Literatur* 1: 147-201.
1879 “Beiträge zur Geschichte der Lautentwicklung und Formenassociation”, *Beiträge zur Geschichte der deutschen Sprache und Literatur* 6: 538-560.

Paul, Hermann – Peter Wiehl – Siegfried Grosse  

Paul, Hermann (ed.)  

Penzl, Herbert  

Picard, Marc  


Pope, Mildred


Prokosch, Eduard


Rooth, E.


Rubach, Jerzy


Russ, Charles (ed.)


Saussure, Ferdinand de


Schane, Sanford


Schane, Sanford A.


Schatz, Josef


Scheer, Tobias


n.d. Von kölnischer Gutturalisierung, verborgener Doppelkonsonanz und Ambiskill. [Ms. Université de Nice.]

Scherrer, Wilhelm

1868 *Zur Geschichte der deutschen Sprache. (1. Aufl.)* Berlin: Duncker.


Schirmunski, Viktor


Schmitt, Alfred


Schrodt, Richard


Schützeichel, Rudolf


Schwarz, Ernst


Stégéral, Philippe – Tobias Scheer


Selkirk, Elisabeth – Jean-Roger Vergnaud


Shyu, Roger W. – Charles-James N. Bailey (eds.)


Sievers, Eduard


Skousen, Royal


Sondergerger, Stefan


Stampe, David


Stark, Marita

1970 *Die sogenannte Rheinische Velarisierung*. [Magisterarbeit, University of Bonn.]

Steriade, Donca


Süreberg, Wilhelm


Szemerényi, Oswald John Louis


Szulc, Alexander

1974 *Diachronische Phonologie und Morphologie des Althochdeutschen*. Warszawa: PWN.
THE MODERN HEBREW ASYNDECTIC RELATIVE CLAUSE: 
THE RISE OF A NEW SYNTACTIC MECHANISM*

Yael Reshef

1. One of the variants of the relative clause in Modern Hebrew is an asynedic construction, which does not include a formal marking of the subordinator by means of a relative subordinator. Instead, it is connected to the main clause by a declined preposition containing a resumptive pronoun referring to the head noun, e.g.

\[(1) \text{ ha-ulan bo neerax ha-diyun} \]

the-hall in-it took-place the-discussion

"the hall in which the discussion took place"

(example from Shlesinger 1994: 76)

This construction is not attested in Hebrew prior to the modern period. It is one of the syntactic constructions that separate it from previous linguistic layers.

The rise of the construction is customarily attributed to the contact between English and Hebrew during the British occupation of Palestine (1917-1948). According to the accepted account, it emerged as a direct calque on the English zero relative: “since in English it is possible, as is well known, to omit the [relative subordinator] (e.g. The family he lived with), one started to do so in Hebrew as well” (Bendavid 1965: 74). The influence of English is considered as the sole factor behind the emergence of the construction.

This article suggests a modification to this explanation. An examination of textual data from the early layer of Modern Hebrew, as well as a structural comparison between the English and the Hebrew constructions, indicate that English had not been the sole factor behind the rise of the construction. The structural comparison reveals significant differences between the languages. The textual data – which were so far overlooked – provide an explanation for these differences and show that change processes started to affect the structure of the Hebrew relative clause long before the contact between English and Hebrew was first created. While English indeed played a role in the rise of the construction, its influence was the culmination of the process rather than its trigger.

A preliminary comment is at place regarding the marked discrepancy between speakers’ perception of the construction and its desired status according to prescriptive normative standards. Among speakers, the construction is considered “most elegant”

* Various points in this article benefited from my consultations with scholars in adjacent fields. While I am grateful to all of them, I would especially like to thank Prof. Benjamin Harashv and Prof. Sally Thomason, whose comments and questions helped me to elaborate and clarify my initial ideas.