SOUND CHANGE IN FUNCTIONAL PHONOLOGY

Paul Boersma
University of Amsterdam, The Netherlands, boersma@fon.let.uva.nl
December 22, 1997

Abstract. *Sound systems may never stop changing, not even if only internal factors are present, because there may always be a better system. Non-teleological random variation of constraint ranking defines a pressure that explains the existence of perpetually rotating cycles of sound changes.*

In a companion paper (Boersma 1997b), I showed that the *symmetries* in inventories of sounds can be described and explained by general properties of motor learning and perceptual categorization, and that the *gaps* in these inventories can be described and explained by asymmetries in articulatory effort or perceptual confusion probabilities. Intimately related with the problem of inventories is the problem of sound change. After all, the inventories have been created in a long series of sound changes, and if inventories seem constructed according to functional principles, these same functional principles must be the driving forces for many sound changes as well.

However, speakers cannot be expected to be able to see ahead to the state that their language will be in after the change: their goal is not to improve the language, but to make themselves understood as quickly, clearly, and easily as possible (Passy 1890). Thus, we have three levels of looking at sound change:

(1) ***The grammar***. In Functional Phonology (Boersma 1997a), the production grammar (i.e., the system that determines the shape of the utterance, given an underlying form) contains constraints that express the drives of maximizing articulatory ease and minimizing perceptual confusion either directly (by disfavouring gestures and favouring faithfulness) or indirectly (by their relative rankings). Strong evidence for the presence of these principles in the production grammar, which handles discrete phonology as well as phonetic implementation, is found in pragmatically based reranking of faithfulness: people are capable of speaking more clearly if they are required to do so. At this level, therefore, goal-oriented processes play a role; in Passy's (1890, p.229) words, "one speaks in order to be understood". For instance, if a language has a voicing contrast in /b/ and /p/, and underlying |b| and |p| are usually pronounced as [b] and [p], a speaker may enhance the contrast by implementing |p| as [pʰ], [p'], or [pː], thus reducing the probability that the listener will perceive the intended |p| as /b/, or by implementing |b| as [β], [ɓ], or [ᵐb].

(2) ***The change***. If many speakers let a certain constraint prevail, e.g., if many speakers implement |p| as [p^h], new learners will include the aspiration in their specifications, thus creating a new underlying segment (= bundle of perceptual features) |p^h|. This change is automatic; it was not a goal of anyone at any time. Though change, therefore, is not *teleological* (there is no *final causation*), it is *functional* in the sense that it is the result of local optimization in the production grammar.

(3) ***The inventory***. As a result of the change, the inventory has improved: a voicing contrast has changed into a voicing-and-aspiration contrast, reducing the average number of confusions. At this level, we can talk in teleological terms again, if only we know what we are doing. This is completely analogous to the common use of teleological jargon in discussions on biological evolution ("why giraffes have long necks"), where everyone realizes that what appears to be a historical gene change is the automatic result of the survival of the fittest (those with the longest necks) in the struggle for life (Darwin 1859), not the result of any goal. However, we saw above that in contrast with genetic change, whose ultimate sources are random mutations, the ultimate sources of (several types of) sound change are directly related to communicative principles in phonetic implementation. In order not to confuse the concrete (1) and abstract (3) uses of teleological terminology, we should refrain from describing change as goal-oriented at the inventory level.

Thus we expect the following types of changes to be frequent:

– *Shifts with conservation of symmetry*. The clearest examples can be read off from small regional variations. Dutch /eː/ is diphthongized to /ei/ in exactly the same regions where /oː/ is diphthongized to /ou/. This may reflect pure diachronic autosegmental behaviour of height contours (like we are used to in the case of tone contours), or result from a quick restoration of symmetry after an initial small imbalance. This restoration may take place as follows. If /eː/ slightly diphthongizes but /oː/ remains the same, listeners have to distinguish two very similar F_1 contours. Quite probably, learners will put these two contours in the same perceptual category and subsequently see no reason to distinguish them in their own productions; the extra F_1 contour has been temporary.

– *Filling of gaps*. If (as the result of a blind sound change) an unnatural gap emerges in a system, subsequent sound changes or lexical selections will hurry to fill that gap. For instance, Latin inherited from Proto-Indo-European the stop system /p t k b d g/ with an unnaturally skewed distribution of voiced stops: though the implementation of voicing would be easiest for the labial plosive, only about 1.2% of all Latin words started with /b/, whereas /d/ accounted for 6% (without the de- and di(s)- words: 2%), and /g/ for 1.5% (as a page count in several dictionaries teaches us). In French, these numbers have become 5%, 6% (2%), and 3%.

A weak interpretation of these facts is that the Proto-Indo-European gap at /b/ was in an unnatural position and that the *local-ranking principle* (Boersma 1997a: §11) caused this gap to be a lexical *accident* in the learner's grammar, thereby allowing Indo-Europeans to freely borrow words with /b/ faithfully (Latin, Greek, Sanskrit) or to fill in the gap with a sound change (Greek: dw → b). A stronger interpretation of the same facts is that French borrowed /b/ to a larger extent than the other voiced plosives; this active de-skewing would presumably involve phonologically-determined choices between synonyms in the lexicon. Though I believe that these choices can be made in the production grammar ("choose the best candidate"), a proof of the controversial factuality of this procedure would require a large empirical investigation.

Whether *natural* gaps, like the lacking /g/ in { p t k b d }, can also be filled, depends on the relative importance of the various factors involved, i.e. it depends on the rankings of the faithfulness, gestural, and categorization constraints (Boersma 1997b).

– Emergence of new gaps. If a system obtains a phoneme at a location where it would be natural to have a gap, subsequent sound changes may create such a gap. Many of the defective stop systems /p t k b d/ used to have a /g/. A "passive" explanation would be that a learner does not hear the difference between /g/ and /k/ as well as the differences in the other pairs, and merges the two. An "active" explanation would be that speakers selectively modify their /g/ so that it becomes perceptually more distinct from /k/. In §1.2, I will show that these active modifications are actually used.

The main idea to be learned from this small typology of functionally explainable sound changes, is that symmetrizing principles ("I have learnt a finite number of types of articulatory gestures and perceptual categories") are just as "functional" as those depending on the biases of the human speech and hearing apparatuses ("minimize articulatory effort and perceptual confusion"). The functional tradition (Passy 1890, Martinet 1955) has always recognized that these principles conflict with each other and that every sound system shows a balance that is the result of a struggle between these principles.

The important question, however, has always been: can and should these principles be expressed directly in a grammar of the language? In Boersma (1997a), I have shown that they can be represented in a production grammar, thanks to the formal constraint-based phonology of Optimality Theory (Prince & Smolensky 1993). In Boersma (1997a; to appear a; 1997b,c; to appear b,c), I argue that a phonological theory based on these principles adequately describes the data of the languages of the world without the need for positing any innate features, representations, or constraints.

After elaborating on the controversy (§1), I will use the remainder of this paper to propose an answer to the irritating question:

> **Q**: "if functional principles optimize sound systems by causing sound change, why do not all languages get better, and why do languages never stop changing?"

The proposed answer will simply be: "because there will always be a better system".

1 Criticisms of functionalism in sound change

Several criticisms have been directed to the unclear definitions and lack of formalizability that used to go with the idea of functionalism. These criticisms come together in Trask's (1996) definitions of both *maximum ease of articulation* and *maximum perceptual separation* as "somewhat ill-defined principle[s] sometimes invoked to account for phonological change". With the gestural and faithfulness constraints of Functional Phonology (Boersma 1997a), however, the principles have received formal definitions that are capable not only of explaining, but also of describing sound patterns.

Apart from the definitions of the principles, the concerted effects and interactions of the functional principles have also met with a poor press. Labov (1994) criticizes the simultaneous functional explanations of *chain shift* as an expression of the preservation of contrast and of *parallel shift* as an expression of rule generalization (i.e., preservation of symmetry):

> "the entire discussion will quickly become vacuous if we lump together explanations based on the facilitation of speech with those that are based on the preservation of meaning." (Labov 1994, p. 551)

However, we can use the constraint-ranking approach of Optimality Theory to combine many explanations without 'lumping'; instead, they are *interleaved*, which makes all the difference. Several functional principles can play a role simultaneously. The existence of parallel chain shifts, by the way, proves that.

1.1 Ohala's "phonetic" approach to sound change

With the definition and formalizability issues out of the way, we can turn to another criticism, directed at the idea that functional explanations invoke goal-orientedness. Ohala (1993) aggressively argues against language change involving "goals":

> "reliance on teleological accounts of sound change is poor scientific strategy. For the same reason that the mature sciences such as physics and chemistry do not explain their phenomena (any more) by saying 'the gods willed it', linguists would be advised not to have the 'speaker's will' as the first explanation for language change". (Ohala 1993, p. 263).

Ohala's own proposal involves synchronic unintended variation, *hypo-correction*, and *hyper-correction*. In his model, normal speech perception involves the process of *correction*, which occurs when a listener restores a phoneme from its contextually influenced realization. For instance, in a language with no contrasting nasality for vowels, the utterance [kɑ̃n] can be reconstructed by the listener as the phoneme sequence |kɑn| that was intended by the speaker, because she knows that every vowel is nasalized before a nasal consonant. Hypo-correction occurs if she fails to restore a phoneme, perhaps because the [n] was not pronounced very clearly, and analyses the utterance as |kɑ̃|.

Hyper-correction refers to the listener restoring a phoneme from the troubled environment although it was not intended by the speaker. Ohala’s example is the Latin change /kwiːŋkwe/ ‘five’ → */kiːŋkwe/. The first [w] in [kwiːŋkwe] may be interpreted by the listener as resulting from the spreading of the second, in which case it would be correct to reconstruct the word as |kiːŋkwe|.

Ohala’s theory accounts for several attested sound changes; for instance, it explains most of Kawasaki’s (1982) data of a general avoidance of /wu/ sequences. However, his anti-teleological position denies the possibility of sound changes (or lexical choices) that seem to preserve the contrast between segments, like (to stay with the */wu/ example) the avoidance of /um/ insertion into /w/-initial stems in Tagalog. In the following section, I will discuss a case that probably does involve contrast enhancement.

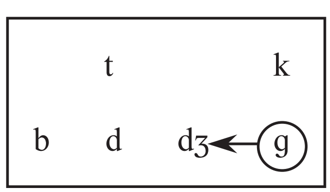
1.2 The story of the fugitive /g/

Ohala (1993) maintains that languages change by misunderstanding of the input, and that goal-oriented drives are never at work. However, we will show in this section that speakers do try to solve the problems that arise when two sounds are hard to distinguish. We will do this by looking at an example that Ohala himself (Ohala & Riordan 1979) has noticed as a tendency in the languages of the world.

Analogously to Ohala’s (1993) account described above, the relative voicelessness of [g], which is due to its short distance to the glottis, would result in misinterpretations of an intended /g/ as /k/. According to Ohala’s reasoning, the only thing /g/ could do, due to the small contrast with /k/, is to be misheard as /k/. If a language used to have [b]-[p] contrasts as well as [g]-[k] contrasts, and now still shows a [b]-[p] contrast but no [g]-[k] contrast, this would have to be due to a coalescence of the velar stops, in particular the conversion of /g/ into /k/.

This would give a merger of /g/ and /k/ in all cases. Surely the /g/ could not travel away from this danger zone, which would be a bad case of teleology? Nevertheless, exactly this is what real languages seem to do: most languages that lost the /g/-/k/ contrast while retaining the /b/-/p/ contrast, did so by converting their /g/ into something perceptually more distinguishable from /k/. Here are a few examples.

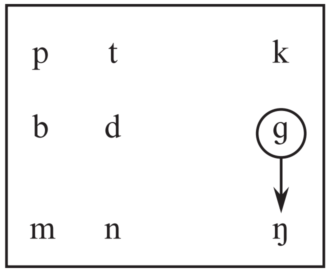
Arabic. In Arabic, an older /g/ was fronted and affricated, and became the palato-alveolar affricate /dʒ/ (Moscati, Spitaler, Ullendorff & Von Soden 1964):



(1)

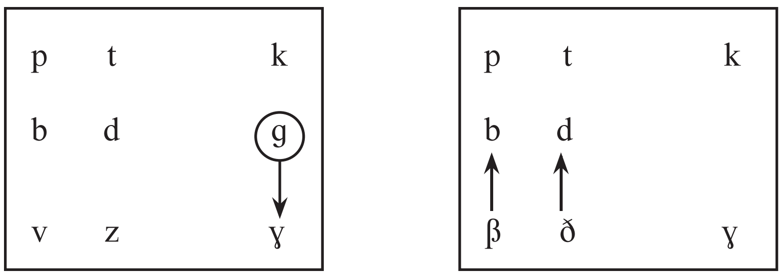
In Arabic, the /g/ problem was solved by creating a new place of articulation.

Japanese. In older standard Japanese, /g/, which acts as a plosive stop throughout the phonology, is pronounced word-internally as the velar nasal stop [ŋ], except in geminates. This makes voicing of the /g/ easier by opening the naso-pharyngeal port:



(2)

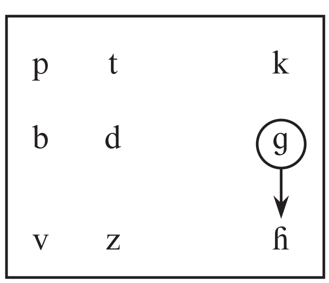
Low German. The Low German /ɣ/ may derive from /g/, which is still heard in some other Germanic languages (left figure):



(3)

The /ɣ/ stayed a voiced velar fricative in most Low Franconian dialects (including Limburgian), though it is /ɦ/ in West Flemish; in most Low Saxon dialects (including Westphalian), it turned voiceless (/x/); in Holland Dutch (including Afrikaans), it became a voiceless pre-uvular fricative /χ/. Many people (Streitberg 1896, Lass 1994) state that the Low German /ɣ/ is the direct descendant of a Proto-Germanic /ɣ/ (right figure above). If this is true, our same functional principle could still explain why original /β/ and /ð/ did become plosives but original /ɣ/ did not, though we would no longer have a counterexample to Ohala's statement here. Note that nearly all Germanic dialects that retained (or developed) a plosive /g/, also have an aspirated /k^h/.

Czech, Slovak, Ukrainian. An original /g/, which still appears in most other Slavic languages, changed into the (usually) voiced "glottal fricative" /ɦ/ (by way of /ɣ/, as still occurs in some southern Russian dialects):



(4)

Andersen (1969) also attributes this change to an enhancement of the k/g contrast, noting that the only Slavic dialects that do not show any signs of change of /g/ (Polish), can be argued to have once possessed an aspirated /kʰ/. Again, we are not sure whether the /g/ is original; actually, things like the Russian rendering of the earlier masculine genitive ending /-ego/ as /-evo/ and, most conspicuously, the fact that final devoicing causes Russian /g/ to become /x/, not /k/, suggest that even the northern Russian "g" was a fricative at the time Russian dropped its final overshort vowels (jers). However, other than in the Germanic case, it is generally assumed that the plosive is original.

Thus, far from merging /g/ with /k/, most languages solved the /g/ problem by replacing their /g/ with a more distinctive sound. Now, is this a bad case of teleology? No, we will see that it is an automatic result of the confusion-reducing principles of phonetic implementation, namely, maximal expression of contrastive perceptual feature values and low faithfulness requirements for non-contrastive feature values. Like the laws of sound change, the biological laws of Darwin are automatic and blind, too, but there, too, we find goal-orientedness on a higher level of abstraction.

1.3 Graduality

The second criticism touches the factual untenability of the requirement for Ohala's model that sound changes cannot be phonetically gradual. This is demonstrably wrong in the case of his own nasality example: languages show all kinds of reproducible sound sequences between [ɑn] and [ɑ̃], as exemplified by the Japanese moraic nasal and by all those languages that replace sequences of vowel + nasal by a nasalized diphthong with an approximant second element; here we find Portuguese ([nɑ̃õ] 'not'), Frisian ([mɛĩskə] 'human being'), and Alemannic ([deĩχu] 'think'; [ɑɰ̃χə] 'butter').

The consistency of graduality and catastrophe is true even within a theory that presupposes universal features. For instance, if the sound /ɔ/ gradually converts into the more close /ɔ̝/, the underlying form may still specify [+open,+mid] or so, and the actual height would be determined at the level of phonetic implementation. However, a learner may reinterpret this sound as a somewhat open /o̞/, resulting in a [–open,+mid] specification. Though the actual pronunciation does not change, the grammar suffered a discrete change, and the new speakers are ready to develop some phonological processes that are appropriate for an underlying /o/. Though a functionalist standpoint could hardly accept this kind of universal feature values, the reshuffling of natural classes after gradual changes is a fact, and so are the discrete changes resulting from reinterpretation of /r/ as /ʀ/ or of /kʷ/ as /p/, which can only be performed by the new language learner.

Below, we will see that small changes will often quickly become large changes, because of a positive-feedback loop: free variation wil cause lowering of faithfulness constraints, and this will increase variation again.

1.4 Merger

The fact that two segments often merge in the history of many languages, is sometimes forwarded as a fact contradicting the functionalist hypothesis. E.g., Labov (1994, p. 551) states that arguments that explain chain shifts with the need to avoid mergers or loss of distinctivity "fail to deal in an accountable way with the fact that mergers are even more common than chain shifts, and that massive mergers do take place, with a concomitant increase in homonymy". I will show why this criticism is not justified. Apart from the fact that articulatory constraints may outrank faithfulness requirements and thus cause merger, there may also be a positive functional drive involved: merger may actually decrease the probability of perceptual confusion.

When populations of speakers who speak different but related dialects come together, the variation in the input to the listener will be high, and the chances of confusions will become high, too. This means that the listeners cannot rely any longer on all the original imported distinctions, and it may become preferable not to take into account any distinction heard between certain sounds.

We can distinguish *discrete* and *gradual* mergers.

New Sittard Limburgian. This is an example of a change in progress: a *discrete* merger, enjoying awareness by the speakers. Original Sittard Limburgian has the reflex of /ôː/ (or, with umlaut, /ø̂ː/) for West-Germanic */au/ before coronals, /h/, and /w/, and word-finally (/ʀôːt/ 'red', /nôːt/ 'need', /ʃôːn/ 'beautiful'), while most neighbouring dialects have the older /ôa/ (/ø̂a/). Besides, it has /ɔ́u/ (with umlaut: /œ́i/) for West-Germanic /oː/ (/ɣɔ́ut/ 'good', /ɣʀœ́in/ 'green'), while the other dialects continue /óː/ (/ǿː/). Both classes have invariable *acute* (falling) accent (Dols 1944) and remain distinct from their *circumflex* (high-mid level) counterparts (/nōːt/ 'nut', /zōːn/ 'son', /ɣɔ̄ut/ 'gold' in all dialects). Comparable relations exist for unrounded front vowels. Now, with the 20th-century mobility around the town of Sittard, the two populations are mixing, and listeners who hear an /ôː/ can hardly use that information to decide between the two word classes. In newer Sittard Limburgian, the contrast between the classes is given up in the production as well: both merge into /ôː/; the imported /ôa/ is considered markedly foreign, and /ɔ́u/ is judged as a distinctive 'real' Sittardism. Given these facts, social factors, like a cooperative effort for indistinguishability, may be involved as well.

If a listener does not take into account a possibly present unreliable distinction, she may actually lower the chances of confusion. Therefore, the typical stages in the process of a *gradual* merger of two segments are:

Stage 1: full contrast. The distinction is psychologically real, is produced in a reproducible manner, and is used for disambiguating utterances.

Stage 2: unreliable contrast. The distinction is psychologically real for people who produce the contrast, but because not everybody does so, they do not use it for disambiguating utterances. In this stage, there are many people who produce the contrast and believe that they can also hear it, though they cannot. The reality of the distinction can only be shown by methods external to the speakers in this stage, for instance by acoustical measurements or a perception test with forced choices (preferably with outsiders for subjects).

Stage 3: near merger. The distinction is not psychologically real and is not used for disambiguating utterances, though some people still produce it in a more or less reproducible manner.

Stage 4: complete merger. The distinction is not produced or perceived any longer.

Thus, *the loss of the perception of the distinction precedes the loss of its production*. Labov (1994) devotes several chapters to this phenomenon. Here are some examples of the intermediate stages of gradual merger.

San Giorgio Resian. Steenwijk (1992) reports a problem in distinguishing a rounded and an unrounded central mid vowel in a Slovenic dialect in the valley of Resia. Informant ML says that there is a distinction (stage 1 or 2), her son LB denies it (stage 3 or 4). A forced-choice classification experiment with ML's speech, with six listeners including four from outside Resia, proved that her three "e"s differed from her three "o"s along the continuum defined by the exaggerated categories "e" and "o", which is compatible with ML being in stage 1 or 2. A forced-choice identification experiment with the same data showed that ML had trouble distinguishing her own utterances, which confines her to stage 2 (LB performed somewhat better; unfortunately, there are no data on his production).

***Dutch* /ɔ/ *and* /o/**. In the larger part of the Netherlands, the distinction between original short /ɔ/ and /o/ is not lexical, but instead the choice of the allophones depends on the fine phonetic structure of the adjacent consonants. More than half of the speakers use both the closed variant before a homorganic cluster of nasal and plosive, and the open variant in words with no nasal or labial consonants. Because nobody knows or hears the difference, it must be stage 3 or 4.

Schouten (1981) performed a number of experiments by which he tried to prove that standard Dutch had a phonemic distinction between the short vowels /ɔ/ and /o/ for speakers from the east of the country (where the distinction occurs uncontroversially in the local Saxon dialects). First, Schouten stated that for every word in a long list, he could tell which of the two phonemes it contained. This psychological reality (stage 1 or 2) was confirmed in a comparison with his brother, who agreed on more than 90% of the words. Nevertheless, Schouten himself could not use the distinction for identification (apparently, stage 2), having been trained for years in *not* hearing the distinction between these sounds. Unfortunately, the experiment stopped here. Rather than doubting the psychological reality of the distinction, Schouten could have proved stage-2 behaviour with a forced-choice classification experiment (say, with exaggerated response categories [a] and [u]) involving listeners from other language backgrounds.

Thus, while mergers do reduce contrasts, they may also reduce confusion. So, while we could defend that the general functional principle of the minimization of confusion is usually implemented as the maximization of perceptual contrast, we see an example of exactly the reverse implementation here.

1.5 Comprehensive approach

Only a comprehensive approach will prove appropriate for the explanation of sound change. Thus, sound change is both preservation of contrast *and* merger, because both can reduce confusion; it is both reduction of confusion *and* facilitation of speech, because both are functional drives; and it is both gradual *and* catastrophic, because acoustic cues determine distinctivity while perceptual categorization determines recognition.

2 An algorithm for function-driven sound change

If we can model sound change, we can also model sound structures by starting from a random set of sounds and letting many sound changes convert this impossible sound system into a natural one. Every sound system is the result of centuries of sound changes. An procedure for arriving at a good sound system could be:

(1) Start out with any vocabulary (e.g., existing or random), and determine its grammar.
(2) *Variation*: propose many different rankings of the constraints in this grammar. Many of these rerankings will propose new sound systems.
(3) *Selection*: choose from the pool of variation the sound system that occurs most often in this pool: a majority decision among the speakers of the language.
(4) Return to step 2.

The criterion for step 3 boils down to a majority decision between competing constraints. We could test this hypothesis by looking at all historical sound changes of any language. Some properties of the criterion are:

- ***Local***: because of the constraint-ranking approach, we do not have to measure effort and confusion in absolute terms; our only concern is whether the satisfaction of articulatory and perceptual constraints improves or not.
- ***Unidirectionality***: if the criterion for step 3 is taken strongly, then if a certain sound change would “improve” the sound system and would therefore be allowed to take

place, the reverse change would not be allowed to take place. This is the part of our hypothesis that makes it in principle testable and falsifiable. A weaker form of our criterion would have a probabilistic interpretation: if a reranking proposes a sound change, this sound change would be possible; sound changes that are more often proposed than others (because they are more common in random variation) have a larger probability of occurrence.

- ***Circularity***: because of the constraint-ranking approach, which cannot assign absolute quality measures to sound systems, it is possible that there does not exist any optimum system and that sound systems keep changing forever. It might be true that the larger the sound system, the smaller the chance that it will ever settle down as an optimal system. In this way, a sequence of several "system-improving" sound changes may eventually result in the re-emergence of the original sound system. A historical example of this circularity is shown in the next sections.

Note that our procedure is a mechanical process without goal orientation, and teleological only at a higher level of abstraction, just like Darwin's survival of the fittest.

3 Changes in an obstruent system

Our example will be a small language with only nine utterances, taken from the "universal" set of voiced, voiceless, and aspirated plosives and voiced and voiceless fricatives articulated with the labial, coronal, and dorsal articulators, followed by [a]. Thus, the utterances of our sample language are drawn from the following set:

ba	pa	p^{h}a	fa	va	
da	ta	t^{h}a	θa	ða	
ɡa	ka	k^{h}a	xa	ɣa	(5)

I shall further assume that the system is symmetric (no more than three manners are used at a time) and stays symmetric across sound changes, and that there are only changes in "manner", not in place. So, if /p/ changes to /p^h/, /t/ and /k/ will change to /t^h/ and /k^h/.

3.1 Hierarchy of articulatory effort

The first hierarchy that we consider, is that of articulatory effort.

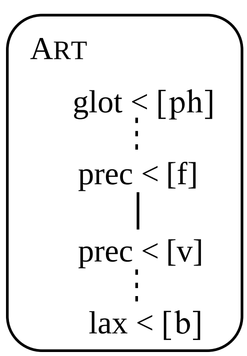
If a complete closing and opening gesture of the tongue tip is superposed on [a] and the larynx muscles are not adjusted, the vocal folds will stop vibrating soon. This is due to the increase in the oral pressure, which causes a rising intraglottal pressure which pushes the vocal folds apart. The resulting sound is a lenis voiceless stop [d̥]. To pronounce a fully voiced [d], the speaker will have to adjust the width of her glottis and may take recourse to several aiding tricks, such as a lowering larynx or slack supraglottal walls. Thus, if we compare /ta/ with /da/, we see that if /ta/ is allowed to be implemented with a passive larynx, it involves one gesture less than /da/ does (though many languages will enhance the contrast with /da/ by means of such tricks as an active widening of the glottis, a raising larynx, or stiff supraglottal walls). That constitutes the first part of the articulatory effort hierarchy: [ta] is easier to pronounce than [da]. This is expressed by the implementation constraint “any laxness gesture should be less effortful than the laxness gesture associated with a typical voiced plosive”, or in short: “lax < [b]” (now taking the somewhat easier [b] as a representative of the three voiced plosives).

The remaining parts of the articulatory-effort hierarchy are the following. Plosives are easier to pronounce than fricatives. A ballistic movement is easier than a controlled movement (it is easier to run into a wall than to suddenly halt one inch in front of it). So [pa] is easier to pronounce than [fa]. This is expressed with the constraint “any articulatory precision should be less than the precision associated with a typical [f]”, or in short: “prec < [f]”.

If /v/ is realized as an approximant, it is spontaneously voiced and therefore does not involve an active gesture to make it voiced, and it requires less precision than /f/ does (it could halt anywhere between one and five inches from the wall). If we allow this freedom for /v/ (in accordance with its use in, say, English or French), the comparison of [f] and [v], only involves a pair of fixed (locally) ranked precision constraints: “prec < [f]” >> “prec < [v]”.

Finally there is [ph]. It involves an abduction of the vocal cords almost to the position during respiration and is therefore more difficult before a vowel than [p], which has no glottal contours at all. We could use the constraint *GESTURE (spread glottis), or its implementation formulation *[aspiration], but sticking with the style of the other articulatory constraints, we will use “glot < [ph]”, which is short for “any amount of glottal widening should be less than the widening associated with a typical [ph]”.

If we assume that making an active glottal opening gesture is more difficult than the precision needed for a continuant, and that this is more difficult again than the implementation of obstruent voicing, we get the following ranking within the articulation (“ART”) family (solid lines denote fixed local rankings, dotted lines denote language-specific rankings):



(6)

With a simplifying move, I shall promote the global effort criterion to the status of a fixed ranking. Thus, I will regard the hierarchy (6) as fixed, though this goes against my cherished local-ranking principle (Boersma 1997a), which would only regard the relative ranking of the two precision constraints as fixed, and all the other rankings as language-specific. In this special case, I justify the fixed global ranking with the idea that global effort measures can predict *tendencies* of asymmetry, and it is these tendencies that will cause some constraints to be on top more often than others in the pool of free variation that we consider the breeding place for sound change.

The hierarchy (6) would prefer [p] as the optimal implementation of an underlying underspecified |labial obstruent + a|:

|labial obstruent + a|	glot < [ph]	prec < [fa]	prec < [v]	lax < [b]
[pha]	*!			
☞ [pa]				
[ba]				*!
[fa]		*!	*	
[va]			*!	

(7)

Later, we will see that such an underspecification is actually a *weak* specification of a non-contrastive feature value. An evaluation as in (7) could occur in a language with a single labial obstruent.

3.2 Hierarchy of perceptual place distinctions

We now assume that the language, next to a series of labials, also has velar consonants with the same manner features. It is likely that voicing obscures place distinctions, so that the /p/-/k/ contrast is larger than the /b/-/g/ contrast, and the /f/-/x/ contrast is larger than the /v/-/ɣ/ contrast. If we take into consideration the commonness of changes of labial fricatives into velar fricatives (Dutch /ɑxtər/ 'behind' < /ɑftər/) and the other way around, we can tentatively propose the following place-distinction hierarchy:

place distances			
	v	40	ɣ
	f	50	x
	b	60	g
	p	70	k
	ph	80	kh

(8)

This hierarchy is the same as the one used by Boersma (1989, 1990) to express the functional principle of maximization of *salience*, which was defined by Kawasaki (1982) as the change in perceptual features as a function of time. Thus, /pha/ has the largest salience, as the two segments /ph/ and /a/ differ in sonorance, voicing, continuancy, and noise. Likewise, /va/ has the lowest salience of the five syllables, because /v/ shares with /a/ its voicing and continuancy features. The hierarchy (8) is also reminiscent of the *sonority hierarchy*.

3.3 Teleological inventory-oriented accounts of distinctivity

On the level of inventory evaluation, we could translate the distinctivity requirement directly into a family of inventory constraints:

Def. *NEAR (*f*: *d*) $\equiv \exists x \in \Im \wedge \exists y \in \Im \wedge x \neq y \wedge \left(\forall g \neq f : g(x) = g(y)\right) \Rightarrow \left|f(x) - f(y)\right| > d$
"If two non-identical segments *x* and *y* in the inventory $\Im$ contrast along the perceptual dimension *f* only, their perceptual distance (along this dimension) is greater than *d*." (9)

Its ranking within a family associated with a specific perceptual tier *f* would be fixed: the lower the distinctivity, the higher the constraint against having the corresponding contrast in the language:

$$\text{*NEAR } (\textit{feature}: d_1) \gg \text{*NEAR } (\textit{feature}: d_2) \Leftrightarrow d_1 < d_2 \qquad (10)$$

In our example (8), the pair /f/ - /x/ violates *NEAR (place: 55) and, a fortiori, the lower ranked *NEAR (place: 65). Let's loosely call these constraints "Δplace > 55" etc.

With evaluation of all pairs, the inventory { ph b v kh g ɣ } turns out to be more contrastive than { p b v k g ɣ }:

	Δplace > 35	Δplace > 45	Δplace > 55	Δplace > 65	Δplace > 75	Δplace > 85
p b v k g ɣ		*	*	**	***!	***
☞ ph b v kh g ɣ		*	*	**	**	***

(11)

In the inventory candidate { p b v k g ɣ }, the constraint "Δplace > 75" is violated three times: by the pair /v/ - /ɣ/, by the pair /b/ - /g/, and by /p/ - /k/. In the other candidate, the last violation is removed, so this inventory shows up as the winner.

This approach is equivalent to the non-OT pairwise inventory evaluation procedure in Boersma (1989, 1990). It also bears a similarity to the OT inventory evaluation approach of Flemming (1995), who, however, only evaluates the *minimal* distance in the system, not all the pairs. For instance, in Flemming's approach, the two systems in (11) would be equally contrastive:

	MINDIST (place) > 35	MINDIST (place) > 45	MINDIST (place) > 55	MINDIST (place) > 75	MINDIST (place) > 85
☞ p b v k g ɣ		*	*	*	*
☞ ph b v kh g ɣ		*	*	*	*

(12)

Even in the inventory candidate { p b v k g ɣ }, Flemming's constraint "MINDIST (place) > 85" is only violated once, because it evaluates only the least distinctive pair /v/ - /ɣ/. In contrast, our all-pairs approach is more in line with the intuitive idea that the need for the enhancement of the /p/ - /k/ contrast (by shifting it to /ph/ - /kh/) should be considered independent from the presence or absence of a /v/ - /ɣ/ pair.

The drawback from this technique is that it suggests that speakers are actively striving for better inventories: both (11) and (12) express teleology at the level of the inventory, a situation that we marked as undesirable in the introduction. Also, this technique suggests that optimization of inventories is a process separated from the OT production grammar. In the next section, I shall replace it with a more accurate and non-teleological formulation based on normal gestural and faithfulness constraints that evaluate the free variation in the production grammar.

3.4 Distinctivity in the production grammar

As phonologists, we all accept the existence of production grammars, i.e. a system that processes lexical information so that it can be pronounced. If the facts of inventories could be described by production grammars, we should not posit the existence of dedicated inventory grammars. As I have shown in Boersma (1997b), inventories can be successfully described with the properties of articulatory, faithfulness, and categorization constraints, all of which are independently required in the production grammar. So, the inventory constraints introduced in §3.3 do not exist, until further evidence forces us to believe in them.

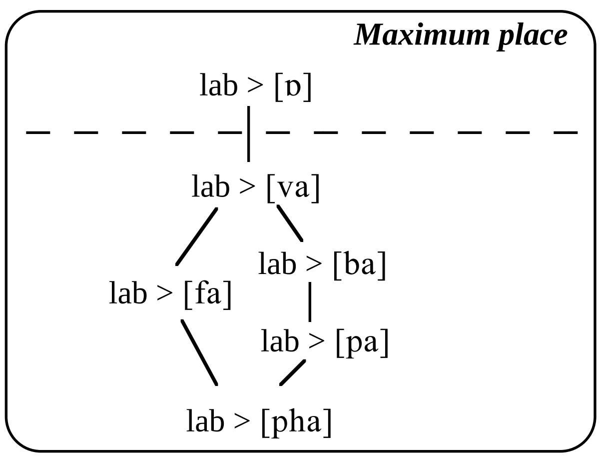
In the inventory formulation, the perceptual-distance constraints evaluated pairs of segments. In a production-grammar formulation, the functional principle of minimization of perceptual confusion has to be expressed indirectly (but locally) in terms of faithfulness constraints and their rankings. For instance, a segment specified as [labial] should be perceived as labial; this may be expressed with the faithfulness constraint PARSE (labial). If the segment contrasts with a dorsal segment, it is desirable that it is pronounced as *very* labial; the more labiality is given to the segment in the process of phonetic implementation, the lower the probability that the listener will mistake it for its dorsal counterpart.

This maximization of unary or peripheral feature values in phonetic implementation is expressed with the following constraint (Boersma 1997a: §8.4):

Def. MAXIMUM $(f\!: v) \equiv \exists x_i \in f_{spec} \wedge \exists y_i \in f_{ac} \Rightarrow \left(x_i = \text{“max”} \Rightarrow y_i > v \right)$
"If a feature x on a tier f is specified as "max", the value of its acoustic correspondent y, if present, should be greater than any finite value v." (13)

Because we cannot assign numeric values to the degrees of labiality needed for the family MAXIMUM (labial), I will abbreviate these constraints in a by now familiar way, giving things like "lab > [ba]", short for "any segment specified as labial should be pronounced with better labiality cues than those present in a typical [ba]". The "any" in this formulation is justified because all labial obstruents in our example are considered to have coronal and dorsal counterparts.

From the distance hierarchy (8), we can infer that labiality cues are best in [pha] and worst in [va]. Therefore, producing labiality cues that are better than those of [pha] is much less important than producing labiality cues better than those found in [va]; this leads to a partial grammar that is parallel to the place hierarchy in (8). Restoring our ignorance of the relative place distinction of the voiceless fricatives with respect to the place distinctions of the non-aspirated plosives, we get



(14)

The location of the somewhat arbitrary constraint in the upper stratum (above the dashed line) expresses the notion of *sufficient* contrast: a segment specified as [labial] should get better labiality cues than the poor ones present in a typical [ɒ]. The lower stratum contains the constraints that express the notion of *maximum* contrast: depending on their ranking with respect to other families, they may be violated or not, as long as the fixed rankings (the drawn lines) are honoured.

We can evaluate labiality directly on an underlying/surface pair, without reference to dorsal obstruents (the ranking values of the labiality constraints, however, do depend on the presence or absence of dorsal obstruents):

|voiceless labial plosive + abava|	lab > [ɒ]	lab > [va]	lab > [b]	lab > [p]	lab > [ph]
pabava		*	**	***!	***
☞ phabava		*	**	**	***

(15)

Like in (7), the underspecification can only occur if |p| is not contrastively specified for [–noise], i.e., if |v| is specified as [–plosive] (or [+continuant]) instead of [+noise]. Our five consonants have the following perceptual features:

	|ph|	|p|	|f|	|b|	|v|
voiced	–	–	–	+	+
plosive	+	+	–	+	–
noise	+	–	+	–	+

(16)

Depending on their contrastive load, some of these specifications are strong and others are weak.

3.5 Poverty of the base

In a production-grammar formulation, we should derive sound changes with the mechanisms of richness of the base and filters, as performed for synchronic descriptions of inventories in Boersma (1997b). Let's start with the obstruent system { p b v }.

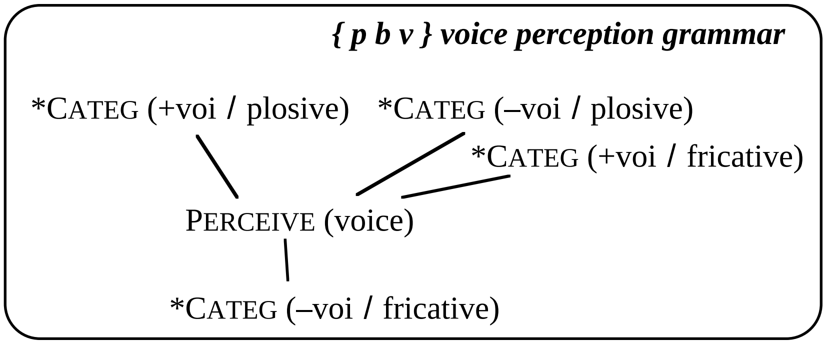
We should first identify a plausible synchronic description of the obstruent system. Apparently, this system uses the two binary-valued perceptual features [voice] and [noise]. If the perceptual categorizations along these two dimensions were independent of one another, the listener would be able to discriminate the four obstruents { p b f v }. In that case, the production grammar should prevent the faithful surfacing of an underlying |f|, presumably realizing it as [v], perceived as /v/. The grammar can achieve this by ranking PARSE (±voice) lower for fricatives than for plosives:

|pabafava|	PARSE (±voice / plosive)	prec < [f]	prec < [v]	PARSE (±voice / fricative)
[pabafava] /pabafava/		*!		
☞ [pabavava] /pabavava/				*

(17)

We see that any underlying |f|, supplied by richness of the base, is realized as [v], so that the surface inventory is { p b v }.

Some listeners may find it disadvantageous to maintain a perceptual voicing contrast in fricatives, only to merge the perceived /f/ and /v/ into a single underlying segment |v| in the recognition phase. These listeners will probably divide the perceptual voicing scale into two values for plosives, and into a single value for fricatives. This can be described by the following perception grammar:



(18)

This grammar will ensure that a pronounced [f] is directly categorized as /v/. Now, the underlying form |pabafava| will always be perceived as /pabavava/, whether |f| is realized as [f] or as [v]. Unless we feel the need for a ghost segment, it is not very useful to maintain an underlying |f| - |v| contrast. In other words, the adaptation of the categorization system to the inventory { p b v } leads to an “impoverished base”, without any occurrences of underlying |f|.

3.6 Hierarchies of manner distinctions

Bringing the manner contrasts into the picture is more complicated, because these cannot be expressed in a single fixed hierarchy like (14): if the perceptual manner space is divided up into three binary features (voice, noise, continuant), we have three separate MAXIMUM (voice), MAXIMUM (noise), and MAXIMUM (continuant) hierarchies, together with three corresponding MINIMUM hierarchies. Moreover, the rankings of these families may depend on the segment involved, i.e. on whether its specifications for these features are contrastive or not. Nevertheless, the rankings *within* these families are fixed. Here are the MAXIMUM (voice) and MINIMUM (voice) families:

Maximum voice

voi (|b|) ≥ [b̥] >> voi (|b|) ≥ [b] >> voi (|b|) ≥ [ɓ]

voi (|p|) ≤ [b̥] >> voi (|p|) ≤ [p] >> voi (|p|) ≤ [ph]

voi (|v|) ≥ [f] >> voi (|v|) ≥ [v] >> voi (|v|) ≥ [w]

voi (|f|) ≤ [v] >> voi (|f|) ≤ [f]

(19)

In the “voi (|b|)” column, we see that for a plosive specified as [+voiced], it is most desirable to be at least as voiced as the voiceless lenis plosive [b̥], but that it is not so important that it is as loudly voiced as the implosive [ɓ]. Likewise, the “voi (|p|)” column expresses the increase in voicelessness as we go along the continuum [b̥] - [p] - [ph]; the “voi (|ph|)” family is similar. Finally, we see that [w] is more voiced than [v], in the sense that [w] is less likely than [v] to be incorrectly categorized as [–voiced].

For the noise feature, the fixed hierarchies are

Maximum noise

noise (|v|) ≥ [v]
|
noise (|v|) ≥ [f]

noise (|b|) ≤ [v]
|
noise (|b|) ≤ [b]

noise (|f|) ≥ [v]   noise (|f|) ≥ [ph]
\ /
noise (|f|) ≥ [f]

noise (|ph|) ≥ [p]
|
noise (|ph|) ≥ [ph]
|
noise (|ph|) ≥ [f]

noise (|p|) ≤ [f]
|
noise (|p|) ≤ [ph]
|
noise (|p|) ≤ [p]

(20)

We know that [f] is more noisy than both [v] and [ph]; these relations account for five of the fixed rankings in (20). The “noise (|f|)” family is not totally ranked, because we cannot (and don’t have to) tell whether [v] or [ph] is the more noisy of the two. Note that the constraints not connected via solid lines in (20) are not a priori ranked with respect to one another.

The two binary features noise and voice do not suffice for distinguishing the five obstruents. Specifically, /ph/ and /f/ have the same representation. So we need a perceptual feature like plosiveness, i.e. the degree to which the surrounding vowels are interrupted by something that resembles silence. If speakers distinguish two values for this feature, they will classify /p/, /b/, and /ph/ as [+plosive], and /f/ and /v/ as [–plosive]. This leads to the the following fixed hierarchies:

Maximum plosive

plosive (|p|) ≥ [v]
|
plosive (|p|) ≥ [f]
|
plosive (|p|) ≥ [b]
|
plosive (|p|) ≥ [p]
|
plosive (|p|) ≥ [ph]

plosive (|f|) ≤ [ph]
|
plosive (|f|) ≤ [p]
|
plosive (|f|) ≤ [b]
|
plosive (|f|) ≤ [f]
|
plosive (|f|) ≤ [v]

(21)

Apart from the subdivision between the plosives and non-plosives, these hierarchies express the idea that [ph] is the strongest and [b] the weakest plosive, and that [v] is even more non-plosive than [f].

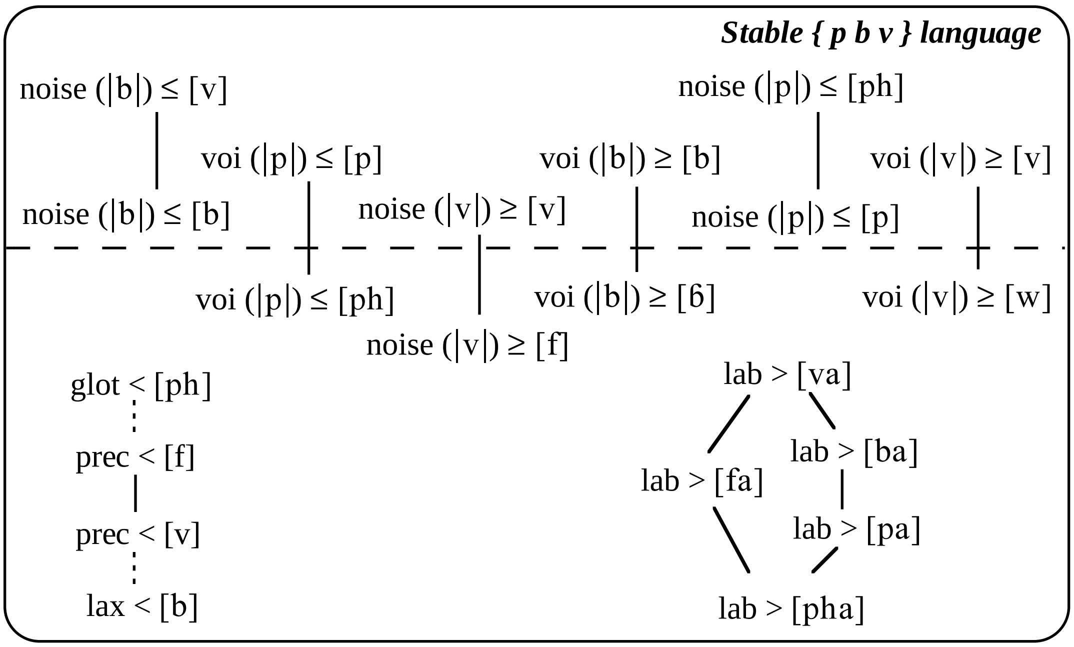
3.7 Stable systems

A stable { p b v } system will have the following specifications:

	p	b	v
[voice]	–	+	+
[noise]	–	–	+

(22)

The six specifications are translated as six families of faithfulness constraints, and the relevant member of each family is undominated:



(23)

Note the seemingly tautological formulation of the six constraints that are just above the dashed line: they are the expressions in phonetic-implementation language of the faithfulness constraints PARSE (+noise), PARSE (–noise), PARSE (+voice), and PARSE (–voice), or if you insist on unary perceptual features: PARSE (noise), FILL (noise), PARSE (voice), and FILL (voice). In (23), we see that the production grammar bundles seemingly discrete phonological constraints with continuous phonetic constraints.

In the grammars, an underlying |v| will always surface as [v], and |p| will surface as [p], independently from the relative rankings of any constraints in the lower stratum (the region below the dashed line). For instance, “noise (|v|) ≥ [f]” may favour the surfacing of |v| as [f], but the voicing-parsing constraint “voi (|v|) ≥ [v]” will prevent that, being in the upper stratum. Likewise, “lab > [pa]” would favour the rendering of |p| as [ph], were it not for the high ranking of the noise-filling constraint “noise (|p|) ≤ [p]”.

3.8 Direct contrastivity and free variation

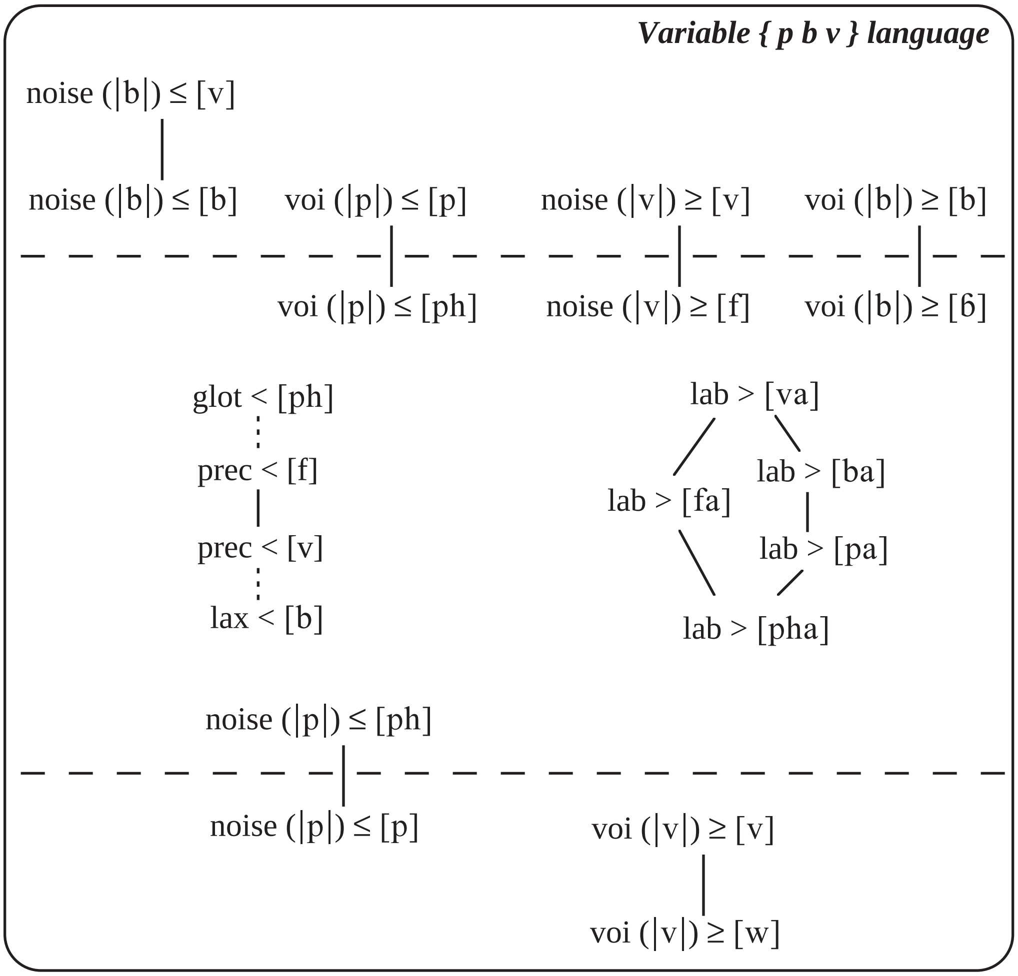
With the impoverished base that results from removing /f/ from the underlying inventory, the path is cleared for underspecified underlying representations. Contrastive specifications for { p b v } can be read from the following table:

	p	b	v
[voice]	–	+	(+)
[noise]	(–)	–	+

(24)

The specifications between parentheses are not directly contrastive, since /p/ has no direct [+noise] counterpart and /v/ has no direct [–voice] counterpart. Several theories connect contrastivity with phonological activity or passivity: in a derivational theory of contrastive underspecification (Steriade 1987), the [+voice] value for /v/ would be filled in at a late stage, probably after any [+voice] values would have been able to trigger assimilation; in a functional constraint-based theory, |v| would be *weakly* specified for [+voice], so that it can easily undergo devoicing, e.g. by assimilation from neighbouring segments more strongly specified for [–voice]. The functional idea behind this connection is that in a system without underlying voiceless fricatives, a |v| can be pronounced as [f] without too many problems with regard to perceptual confusion.

Let us translate the contrastivity argument into a constraint ranking. The voicing specifications for |p| and |b| are expressed as *REPLACE (voice: –, + / plosive) and *REPLACE (voice: +, – / plosive), or more loosely as PARSE (–voice / plosive) and PARSE (+voice / plosive), or more readably as the specification constraints |p| → [–voi] and |b| → [+voi] or their continuous counterparts like “voi (|p|) ≤ [p]”and “voi (|b|) ≥ [b]”. The strength of these constraints is expressed by putting them on top of the grammar. Likewise, the weakness of the specifications |v| → [+voi] and |p| → [–noi] is expressed by putting their continuous counterparts “voi (|v|) ≥ [v]” and “noise (|p|) ≤ [p]” at the bottom. The remaining constraints are ranked in an intermediate stratum:



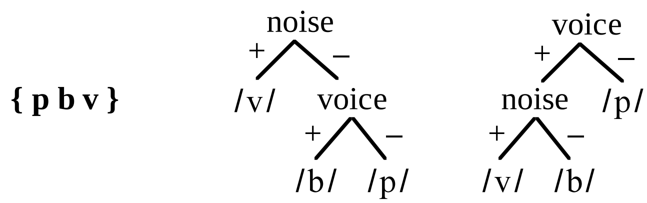
(25)

Note that the six “tautological” constraints of the faithful grammar (23) are now located either at the bottom of the highest or at the top of the lowest stratum. These six, therefore, express the principle of *sufficient contrast*; their lower-ranked fellows are in the lower strata, expressing the universally less important principle of *maximum contrast*.

If the rankings of the intermediate constraints are varied randomly, the division between the contrastive specifications at the top and the redundant specifications at the bottom will lead to a certain amount of free variation between the implementations [f] and [v] for |v|, and [p] and [ph] for |p|. Note that these alternations are exactly those that allow the speaker to be understood, i.e. those that allow the listener to reconstruct the underlying form as long as she either perceives [f] directly as /v/ or successfully reconstructs a perceived /f/ as |v| (and analogously for [ph] and |p|).

3.9 Indirect contrastivity and free variation

The concept of direct contrastivity in (24) yields an amount of underspecification too large for the purpose of determining the allowed degree of free variation: if /p/ is specified solely as [–voice] and /v/ is just [+noise], then there is no reliable contrast between these two segments. It would seem that we cannot freely vary the redundant feature values, as that would allow both /p/ and /v/ to be realized as [f]. So we should keep either the [+voice] specification of /v/ or the [–noise] specification of /p/, depending on what we consider to be the more basic feature. Both these possibilities can be represented by simple *feature trees* (in support of Jakobson, Cherry & Halle 1953, but contra Frisch, Broe & Pierrehumbert 1997), with segments as leaves:



(26)

The left-hand representation shows that /v/ can contrast with /b/ and /p/ at the same time. In such a situation, /v/ is allowed to be pronounced voiceless, and /p/ must stay noiseless. This is a reasonable description of the situation in Dutch (disregarding the somewhat marked lexically voiceless fricatives), where voiced fricatives are only weakly specified for [+voice] (they devoice after any obstruent).

The feature-tree underspecification strategy is not only useful for predicting variation, but would also solve a technical problem (noted by Archangeli 1988) with an algorithm for contrastive underspecification that marks only those feature values whose reversal would result in a different segment in the inventory. For instance, the Dutch short-vowel system might be represented with contrastive underspecification as

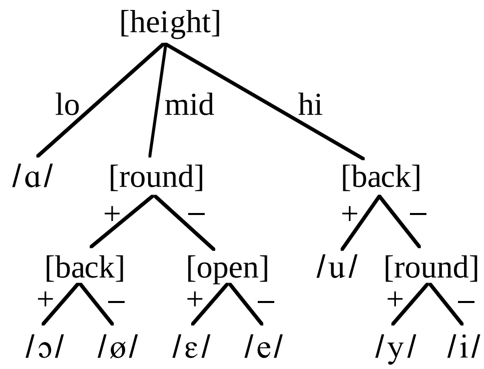
	i	y	u	e	ø	ɛ	ɔ	ɑ
[open]	(–)	(–)	(–)	–	(–)	+	(+)	(+)
[mid]	–	–	(–)	+	+	(+)	(+)	(–)
[round]	–	+	(+)	–	+	(–)	(+)	(–)
[back]	(–)	–	+	(–)	(–)	(–)	(+)	(+)

(27)

The specifications between parentheses are not directly contrastive. The procedure leaves both /ɔ/ and /ɑ/ without any specifications at all. Surely these segments do contrast! This problem cannot be resolved by using different height features (unless /ɑ/ is changed to [+mid]; but the problem is in the algorithm, not in the data).

To solve the ambiguities in (27), we will have to mark some feature values as contrastive. We have to decide which is the most basic feature; for /ɔ/ and /ɑ/, this is perhaps the height distinction, but this will depend on the language. The solution can be graphically represented by feature trees, which guarantee exactly the right amount of specification. Several feature trees are possible for the Dutch data, but I will choose the one that I think most closely reflects the behaviour of the Dutch speaker/listener:

Dutch short vowels



(28)

It may well be that this tree tells us more about the psychological realities of Dutch vowels than any feature matrix. There are seven underspecifications, most of which are reflected in regional or positional variations:

(1) The underspecifications of the rounding of /u/ and the backness of /ɛ/ and /e/ do not contribute to variation, because the gesture for non-low back unrounded vowels (styloglossus plus orbicularis oris activity) does not belong to the Dutch repertory of articulations (any unrounded back /ɑ/ is implemented without styloglossus). For the high vowels, I could have exchanged the [back] and [round] nodes; this would have given an /i/ underspecified for backness.

(2) The underspecification of the backness and rounding of /ɑ/ corresponds to the high level of regional variation of this sound. Regarding the fact that the lip shape is quite different from that of the long low vowel /aː/, we may call its usual pronunciation rounded. Fronted realizations also are common. Specifying a single low vowel for no other than height features relieves us of the need of arbitrarily assigning it a [+back] or any other specification.

(3) The underspecification of the openness of /ɔ/ corresponds to the high degree of regional and positional variation of this sound. For instance, the pronunciations of Dutch *hondehok* ‘dog kennel’ are (from most to least common): [ɦondəɦɔk], [ɦondəɦok], and [ɦɔndəɦɔk]; any pronunciations with heights between [o] and [ɔ] are also perfectly normal. Underspecifying this vowel for openness allows us not to try to identify its height artificially with that of /ɛ/ or /e/.

(4) The underspecification of the openness of /ø/ corresponds to its special relation with the [œ] sound. Dutch listeners tend to hear foreign [œ] sounds, as in German [gœtɐ] ‘gods’ or French [œf] ‘egg’, as a Dutch /ø/. Even many speech scientists transcribe Dutch /ø/ as “œ”, though it is phonetically the rounded counterpart to /e/ (same F_1, same tongue-fronting gesture), so that we would expect the transcription “ʏ”, analogously to the traditional transcription “ɪ” for /e/. We may further note that /ø/ does not pattern with /e/ at all; instead, there are some words in which /ø/ alternates with /ɔ/ (/ʋɔʀɣə/ = /ʋøʀɣə/ ‘strangle’, etc.), so that the backness correlation between these sounds suggested in (28) is not strange at all.

We can now state the relation between contrastivity and free variation precisely:

The maximum free variation.

> "Segments are allowed to vary freely as long as the listener can easily reconstruct the underlying form. The largest allowed amount of free variation is achieved with random reranking of intermediate constraints, keeping directly or indirectly contrastive specifications fixed at the top and redundant specifications fixed at the bottom." (29)

Our modification of the notion of contrast leaves us with a { p b v } system where only the specification of /v/ voicing ranks in the lowest stratum.

3.10 Where { p b v } will go

Suppose that a language goes from the stable { p b v } system (23) to the (modification of) system (25), which allows some variation of voicing in |v|. In a stable system, the relative rankings of the constraints in the lower stratum cannot be determined, so we may safely suppose that speakers have some invisible random variation here. When its non-contrastivity causes the voicing specification for |v| to drop below the formerly lower stratum, as in (25), the rankings in the emerging intermediate stratum become exposed: they determine whether an underlying |v| is pronounced as [v] or as [f].

The variation in (25) is subject to the local-ranking principle, which ensures that e.g. "lab > [ba]" will always be ranked above "lab > [pa]". Constraint pairs that are not directly or transitively locally ranked, can be ranked in a speaker-specific manner. Whether an underlying |v| surfaces as [v] or [f], there are two constraints that are violated in either case: "prec < [v]" and "lab > [fa]"; and there is one constraint that is satisfied in either case: "noise (|v|) ≥ [v]". For the voicing of an underlying |v|, there remain thus four relevant constraints: "noise (|v|) ≥ [f]", "prec < [f]", "lab > [va]", and "voi (|v|) ≥ [v]". Because of the redundancy of the voicing specification of |v|, the last of these is always ranked at the bottom, so the outcome of an underlying |pabava| will depend on the relative rankings of the remaining constraints. If "noise (|v|) ≥ [f]" is ranked on top, the noise contrast will be enhanced:

|pabava|	noise (|v|) ≥ [f]	prec < [f]	lab > [va]	voi (|v|) ≥ [v]
[pabava]	*!		*	
☞ [pabafa]		*		*

(30)

If "prec < [f]" is on top, the input surfaces faithfully:

|pabava|	prec < [f]	noise (|v|) ≥ [f]	lab > [va]	voi (|v|) ≥ [v]
☞ [pabava]		*	*	
[pabafa]	*!			*

(31)

And if “lab > [va]” is on top, the place contrast is enhanced:

|pabava|	lab > [va]	noise (|v|) ≥ [f]	prec < [f]	voi (|v|) ≥ [v]
[pabava]	*!	*		
☞ [pabafa]			*	*

(32)

If the stable system had a randomly ranked lower stratum, the variable system will have a randomly ranked intermediate stratum. We may guess, therefore, that the three rankings discussed here occur with equal frequency among the speakers. Thus, the majority of speakers with low-ranked |v| voicing will pronounce an underlying |v| as [f]. Underlying |pabava| now often surfaces as [pabafa], which is either directly perceived as /pabava/ or easily reconstructed as |pabava|.

3.11 Graduality and catastrophe

If a specification becomes a little bit redundant, it is going to fall down the constraint hierarchy by a little amount. This will cause the exposure of some formerly invisible randomly ranked constraints, and this will cause some free variation. This free variation will lead the listener into relying less on the presence or absence of the weakened specification. This will invite speakers (at least those who are listeners as well) to lower the faithfulness specifications again. This positive feedback loop may cause the original little sound shift to become large.

Returning to our { p b v } example: after the variation has caused low-ranked |v| voicing in all speakers, the majority of speakers will say [f]. If the implementation rule |v| → [f] becomes general, the next generation will almost certainly perceive the result as /f/ and see no reason to analyse it as anything but a fricative |f| that is underlyingly (weakly) specified as [–voice]. We then have a genuine { p b f } system.

Thus, to arrive at the preferred directions of sound change, we have to assess the hierarchies of articulatory effort and perceptual distinctiveness. We do this by randomly varying the ranking of the three constraint families that handle effort, place perception, and manner perception, though still keeping them internally contiguous.

Thus, the three constraint families together favour the inventory change { p b v } → { p b f }. In the same way, we can predict a lot of other sound changes.

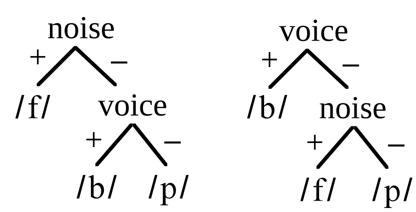
3.12 Unidirectionality

After the { p b v } system has become { p b f }, it will not return to { p b v }. This is because the representation of a { p b f } system that would show the free variation needed for going to { p b v }, would look exactly like the representation of a { p b v } system that would show the free variation needed for going to { p b f }, namely, the feature tree on the left-hand side of (26); and the relevant variation grammar is analogous to (25), with ranking of “voi (|f|) ≤ [f]” in the bottom stratum. The surface variation, therefore, will be the same as that discussed in §3.10, with a majority of speakers pronouncing an underlying |f| as [f]. Learners will specify this sound weakly, but as voiceless, not voiced. If the system ever becomes *stable*, in the sense of (23), it will be with a strongly voiceless |f|.

3.13 Why { p b f } is such a good system

The { p b f } system is by far the most common three-obstruent system: in Maddieson (1984), it occurs twice as often as the second most common system, which is { ph p b } (Boersma 1989:122). With the same features as in (25), its possible feature trees are

{ p b f }



(33)

As we saw in §3.12, the left-hand representation will not lead to sound change. I will now show that the right-hand representation is also optimal within the pool of variation that it allows. In the relevant grammar, /b/ is weakly specified for [–noise], so its realization is allowed to vacillate between [b] and [v] (a situation reminiscent of the Spanish obstruent system). The relevant constraints in the intermediate stratum (i.e., those that give different evaluations for [b] and [v]) are “voi (|b|) ≥ [v]”, “prec < [v]”, and “lab > [va]”. The following three tableaux evaluate the two candidates:

|pabafa|	voi (|b|) ≥ [v]	prec < [v]	lab > [va]	noi (|b|) ≤ [b]
[pabafa]	*!	*		
☞ [pavafa]		**	*	*

(34)

|pabafa|	prec < [v]	voi (|b|) ≥ [v]	lab > [va]	noi (|b|) ≤ [b]
☞ [pabafa]	*	*		
[pavafa]	**!		*	*

(35)

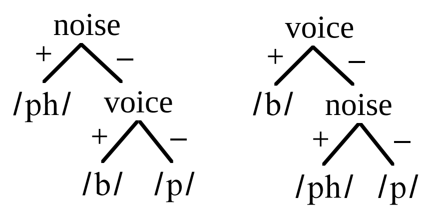
|pabafa|	lab > [va]	voi (|b|) ≥ [v]	prec < [v]	noi (|b|) ≤ [b]
☞ [pabafa]		*!	*	
[pavafa]	*!		**	*

(36)

Thus, the enhancement of the voicing feature is the sole supporter for [v]. The other two drives (enhancement of labiality and minimization of precision) prefer [f]. Hence, the system { p b f } will not change.

From the two representations in (33), we must conclude that the { p b f } system is a *sink* for sound change: both the { p b v } and { p v f } systems can be shifted to { p b f }, and the reverse shifts are impossible (with these two features). And there is a third system that may turn into our optimal system: the all-plosive system { ph p b }. Its feature systems can be described as

{ ph p b }



(37)

Plosiveness is not contrastive in this system, so *in both grammars* [ph] may alternate with [f]. There are three relevant intermediate constraints again. For an underlying |phabapa| with weak "plosive (|ph|) ≥ [ph]", the faithful candidate [phabapa] is preferred only by "lab > [pa]". The candidate [fabapa] wins, because it is supported by the other two, namely "noi (|ph|) ≥ [f]" and "glot < [ph]", which is ranked higher than "prec < [f]". This spirantization of the aspirates at all three articulators in an all-plosive system has probably occurred in Proto-Latin (Classical Greek /phero:/ 'I carry' versus Latin /fero:/). Look ahead to figure (44).

3.14 Tunnelling in Greek

When we compare Classical Greek with modern Greek, we see that an original { ph p b } system became { f p v }. There seem to have been two possibilities, given that the feature-tree representations have seemed to allow only single changes:

- The sequence { ph p b } → { f p b } → { f p v }. As we have seen, our model does not allow the second step.
- The sequence { ph p b } → { ph p v } → { f p v }. The first step is not allowed: it would involve enhancement of the [+voice] specification of |b| at the cost of its redundant specification for [–noise], but it would fail because two constraint families militate against it: maximum labiality ("lab > [ba]") and maximum ease (the allegedly fixed global ranking "prec < [v]" >> "lax < [b]").

But note that if |b| is weakly specified for [–noise], |ph| can still be realized as [f]. This is because /ph/ and /f/ have the same perceptual noise and voice values; their values for [plosive] are different, but this feature is not contrastive in { ph p b }.

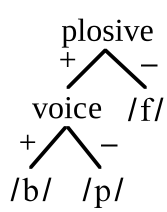
So [fapava] is actually a serious candidate for the implementation of |phapaba|. From the discussion above, we see that three relevant intermediate constraints favour the faithful [phapaba] candidate over [fapava]: maximum labiality in |ph| as well as in |b|, and "prec < [v]". Also three constraints favour [fapava]: maximum voicing in |b|, maximum noise in |ph|, and "glot < [ph]". So it looks a tie. But the constraint "prec < [v]" can never vote in favour of [phapaba], because it can never be ranked higher than "glot < [ph]" if our global fixation of the relation (6) between these constraints holds. Therefore, the articulatory constraints always favour [fapava], the labiality constraints favour [phapaba], and the enhancement constraints favour [fapava], which wins.

It seems, now, that the combined, "tunnelling", change { ph p b } → { f p v } is possible. It requires the two Greek spirantization processes to have been simultaneous. According to Sihler (1995), they should both be dated around the first century A.D.

3.15 Why { p b f } is not the best system

The most common inventory { p b f } is good, but { ph b f } proves to be better if the inventory is represented with the help of the [plosive] feature:

{ p b f }



(38)

Because of the choice of features (perceptual categorization), all [noise] specifications are absent, allowing |p| to surface as [ph]. This candidate wins on voicing contrast and on labiality, at the cost of implementing aspiration.

Tree (38) also shows that |f| has a weak specification for [–voice], suggesting that |pabafa| could be rendered as [pabava]. In (30)-(32), we have already compared these two candidates, but within a different feature set. The evaluations (31) and (32) can more or less be copied for the current situation, but (30) becomes

|pabafa|	plosive (|f|) ≤ [v]	prec < [f]	lab > [va]	voi (|f|) ≤ [f]
☞ [pabava]			*	
[pabafa]	*!	*		*

(39)

This is the small difference between noisiness and continuancy: [f] is more noisy than [v], but [v] is more continuant than [f]. Suddenly, { p b f } can change into { p b v }; it must have happened at some time in the history of Dutch.

The combined change of { p b f } directly into { ph b v } also seems possible, because (38) shows a combination of a weak voice specification for |f| and an absence of any specifications for noise. The aspiration of |p|, therefore, only enhances the voicing contrast with |b|, without diminishing any noise contrast, since listeners hear continuants instead of fricatives; the continuancy contrast is also enhanced. But the articulatory constraints disfavour the change: the precision gain in { ph b v } can never outweigh the aspiration loss, according to (6). Fatally, the labiality constraints also disfavour it: the gain in [p] → [ph] can never outweigh the loss in [f] → [v], according to (14).

So only the reverse combined change { ph b v } → { p b f } is allowed. But still, the inventory { p b f } can become { ph b v } in two steps. First, it can become { ph b f }, as we have seen. After that, { ph b f } can become { ph b v }, within the feature representation (38), since precision and continuancy improve, at the cost of labiality.

We now have a mini-cycle { p b f } → { ph b f } → { ph b v } → { p b f }. In contrast with the even shorter cycle { p b f } → { p b v } → { p b f } that we saw above, this 3-cycle works within a single feature representation (38), so it may well occur in a single language within a relatively short period (assuming that a feature tree tells us something about the grammar, so that we know that most speakers learn the same tree). We find another 3-cycle within the same inventory representation when we realize that { p b v } can become { ph b v }, for the same reasons as { p b f } → { ph b f }. Widespread regional variation between the four inventories compatible with the feature tree (36) (i.e., a voiceless or aspirated plosive, a voiced plosive, and a fricative), is found in the dialect continua of the Netherlands, Northern Germany, and Britain.

Since the aspiration of a voiceless plosive as a means for enhancing the perceptual contrast with its voiced counterpart is most urgent for dorsals, we can derive the following hypothesis of a correlation between aspiration and dorsality:

Suspicion of the dependence of aspiration on a velar voicing contrast:

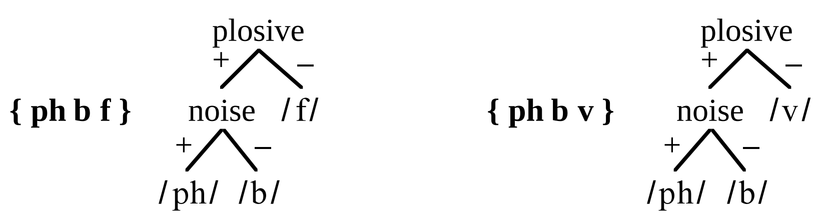
> “in languages that have a /g/ and a /k/, the /k/ may become aspirated (of course, if there is not already a /k^h/). The /p/ and /t/ may then, and only then, also become aspirated. Therefore, we do not expect many aspirates in languages with a gap at /g/.” (40)

This hypothesis, which we saw to be able to account for the Germanic and Slavic distributions, could be checked by the data of more of the world's languages, as they have been compiled in Maddieson (1984), which contains information on the sounds of 317 typologically evenly distributed languages. Of these languages, 88 have a coronal and velar stop system of { t d k g } and 13 have { t^h d k^h g }. However, the languages with a gap at /g/ show a slightly different pattern: 10 languages have a coronal and velar stop system of { t d k } whereas no language has { t^h d k^h }. So, of the languages with /g/, 13% has aspirates, and of the languages with a gap at /g/, 0% has aspirates. Alas, there are not enough data statistically to confirm the dependence of aspiration on the presence of a velar voicing contrast (a Fisher exact test gives $p = 0.27$).

3.16 The second Germanic consonant shift

The proposed change { ph b v } → { p b f } feels somewhat strange: it combines two heterogenous sound changes, in contrast with the Greek double spirantization. Furthermore, its evaluation hinged on a certain way of counting how many of the six relevant constraint rankings spoke for and against it: we could have said that it only got the articulatory and labiality votes, and that two faithfulness constraints opposed it; this is in sharp contrast with the Greek { ph b p } → { f v p } change: no matter how we count votes, it is always favoured. So let us concede that inventories that adhere to (38) will eventually get an aspirated plosive.

I will now show that these { ph b f } and { ph b v } systems favour devoicing of their /b/ segments. First, we can immediately see that "lab > [ba]" and "lax < [b]" will always favour the [b] → [p] change. So we only have to show that this change does not violate any undominated faithfulness constraints, i.e., that we can find feature trees that show weak voicing specifications for /b/. Here they are:



(41)

In both trees, |b| is unpecified for voice, since the voicing feature does not belong to the listener's repertoire of perceptual dimensions. Therefore, speakers are free to vary their voicing of this sound, and the articulatory and place-distinction drives will force its voiceless realization, despite the resulting proximity between |ph| and the new (reanalysed) |p| segment. This devoicing of originally voiced plosives is known in various degrees from Old High German, present-day Alemannic, English, Danish, and Icelandic; the result is typically "lenis" (weak and short) voiceless, and this sound contrasts with an aspirated or "fortis" (longer and stronger) voiceless sound.

The resulting inventories are { ph p f } and { ph p v }. Still within the representation of (39), the first of these can change into the second, because that would improve continuancy (the [–plosive] specification) of |f| and decrease the required precision, at the cost of a loss in labiality cues.

We are left with a single { ph p v } inventory. We met it before in the discussion of Greek: with a voice-noise tree like the one in the right-hand side of (37), its |v| is weakly specified for [+voice], and it will become pronounced as [b] because maximum labiality and ease numerically outweigh its [+voice] specification.

We have now found the five-cycle { p b f } → { ph b f } → { ph p f } → { ph p v } → { ph p b } → { f p b }. Note that the lexicon is not back where we started: the three obstruents have rotated.

The spirantization of |ph| could have occurred directly in { ph p v }, again with the voice-noise features (37): the noise contrast between |ph| and |p| is enhanced (note that |v| was only weakly specified for [+noise]), and the loss of aspiration is an articulatory gain (labiality cues deteriorate). The result is the { f p v } inventory that we met before, and we have identified another cycle.

The presence of a [pf] or [f] is what distinguishes High German from the other Germanic languages, which have not come further than [ph].

Here is a summary of all the allowed sound changes discussed above:

From:	**To:**	**Feature tree:**	**Constraints for:**	**Constraints against:**
p b v	p b f	noi voi	+noi, lab	prec-f
	ph b v	plos voi	–voi, lab	glot
p b f	ph b f	plos voi	–voi, lab	glot
	p b v	plos voi	–plos, prec-f	lab
p v f	p b f	voi noi	prec-v, lab	+voi
ph p b	f p b	voi noi, noi voi	+noi, glot	lab
	f p v	voi noi	+noi & +voi, glot	lab
ph p v	ph p b	voi noi	prec-v, lab	+voi
	f p v	voi noi	+noi, glot	lab
ph b v	p b f	plos voi	lab, glot	–plos & –voi
	ph p v	plos noi	lab, lax	
ph b f	ph b v	plos voi	–plos, prec-f	lab
	ph p f	plos noi	lab, lax	
ph p f	ph p v	plos noi	–plos, prec-f	lab

(42)

All these cases are also visible in (44).

4 The Germanic consonant shifts

Let us now set out to explain the circularity of the Germanic consonant shifts in word-initial position.

4.1 Data

In the first Germanic sound shift (between Proto-Indo-European and Old Germanic), original voiceless stops became voiceless fricatives, original voiced stops became unvoiced, and original 'murmured' stops became plain voiced stops (or voiced fricatives). The second Germanic consonant shift (from Old Germanic to High German) more or less repeated the first. Here is a simplified review of these historical sound changes (cognate Latin words have been added for comparison):

Proto-IE	*p	(*b)	*bɦ	*t	*d	*dɦ	*k	*g	*gɦ
Latin	pel		fol	tri	dw-	(fa-)	kan	gel	(host)
Old Germanic	fel	poːl	bal	θri	tw-	doː-	xund	kald	gast
High German	fel	pfuol	pal	(dri)	tsw-	tuo-	(hund)	kxalt	kast
gloss	'skin'	'pool'	'leaf/ ball'	'three'	'two'	'do'	'dog'	'ice/ cold'	'guest'

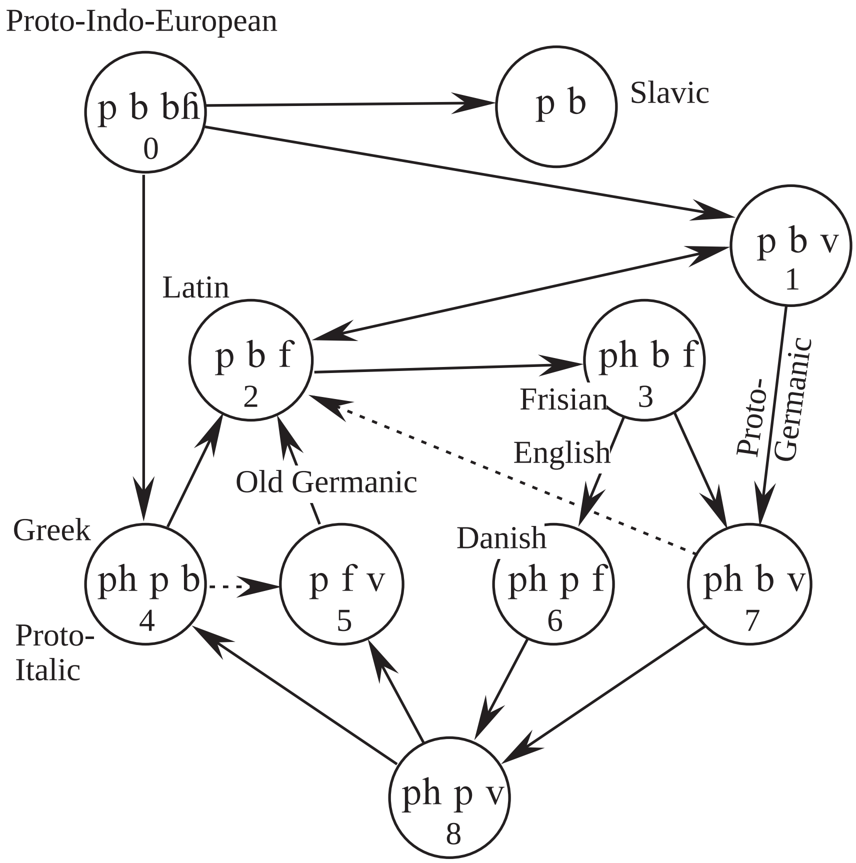
(43)

The endings of the words have been suppressed. Between parentheses, we find the results of some changes that disturb the general pattern:

- The Latin change from the PIE system { bɦ dɦ gɦ } through word-initial { ph th kh } and the non-strident fricatives { ɸ θ x } to the classical { f f h }, with merger of the two anterior fricatives.
- The Common Germanic change of /x/ into /h/.
- The German change of /ð/ into /d/.

4.2 The preferred direction of change

The eight three-consonant systems discussed in §3.16 are all shown in figure (44). The arrows represent the directions of sound change that are preferred according to §3. These preferred directions are equal to the directions of the first and second Germanic consonant shifts in word-initial position. The solid arrows denote the changes of a single segment, the dotted arrows those of two segments. The two possible systems that are not shown are {b, v, f} and {ph, f, v}. These would only have arrows that point away from them.



(44)

4.3 Proto-Indo-European obstruents

There has been some debate about the nature of the plosives in Proto-Indo-European. Gamkrelidze & Ivanov (1973) considered the system { p bɦ, t d dɦ, k g gɦ } typologically unnatural, because in most languages with voiced aspirated plosives, there are also voiceless aspirated plosives. Moreover, there was a gap in the labial system, which lacked the plain voiced plosive [b]. This is a typologically marked situation, too: if a language lacks one of the voiced stops { b d g }, it is usually [g] that is missing (for good reasons, as we saw). However, from a system of ejectives { p’ t’ k’ }, the first element to drop out is usually the labial [p’] (for equally good reasons). Therefore, Gamkrelidze & Ivanov’s proposal included ejectives instead of plain voiced stops, and the voiceless plosives would have been aspirated, because the voiced aspirates would otherwise have no unmarked (voiceless) counterparts. The system resulting from this theory is thus { ph bɦ, th t’ dɦ, kh k’ gɦ }. A related theory (Beekes 1990) does have plain voiceless plosives, but the aspirates are voiceless: { p ph, t t’ th, k k’ kh }.

These ‘glottalic’ theories may be typologically and phonetically more satisfactory than the traditional account, but they run into severe problems if we try to find the paths that led from the PIE system to the systems attested in the daughter languages. The theory by Beekes poses the largest problem: the aspirated stops would have become fricatives in Latin, voiced aspirates in Sanskrit, and voiced stops in Slavic and Germanic. By contrast, voiced aspirates are obviously problematic in many sound systems, so it would only be natural that the problem has been solved in so many ways in the various branches. As we have seen in our discussion on the fate of [g], problematic consonants are struggling to find a different place in the system, and they do so in any suitable way. So, while [g] had the choice to develop into [ɣ], [ŋ], [ɦ], or [dʒ], the breathy stop [dɦ] had the choice of changing into any neighbouring consonant, be it [th] (Greek and Latin), [d] (Slavic), or [ð] (Germanic and Latin). Thus, the diversity of the reflexes in the daughter languages points to a problematic ancestor. However, if we make the unconventional move of regarding the Slavic data as evidence of an earlier stage with voiced fricatives (as in Germanic, see §1.2), the sequence th → θ → ð → d may be defensible.

Secondly (in both glottalic theories), the ejectives would have become plain voiced stops in Italic, Greek, Celtic, and Slavic, independently. The classical theory has far more straightforward reflexes of its plain voiced stops: they stayed that way in most branches, but became voiceless in Germanic, a natural process which repeated itself later in High German.

Can we have both? Yes. The glottalic theory may well be the result of an internal reconstruction on Proto-Indo-European, and the change of { th t’ dɦ } or { t t’ tɦ } into { t d dɦ } may well have occurred before the split.

The top left of figure (44) thus shows the PIE system ‘0’, *with* the notoriously problematic /bɦ/. In Old Greek, for instance, this sound changed into /ph/, giving system ‘4’.

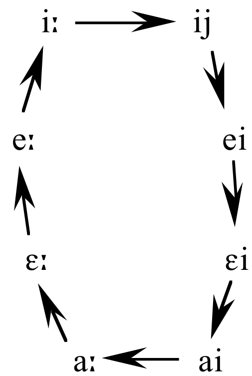
4.4 Circularity

In (44), we see several cycles of optimization, as discussed in §3.15 and §3.16.

If we assume a Pre-Germanic system like { p b β } (numbered ‘1’ in figure (44); we ignore the difference between /β/ and /v/), we see that there are two different routes to arrive at the favoured sound system { p b f }. Most authors state that Latin took the same route as Greek (Sihler 1995), at least in initial position (Stuart-Smith 1995). The second route to the preferred { p b f } starts with aspirating the /p/, which is what Proto-Germanic may have done. This language then necessarily obtained system ‘8’, then ‘4’ or ‘5’, and then number ‘2’. This is why Latin and Old-Germanic both feature the favoured { p, b, f } system, though these systems are shifted relatively to one another.

5 Vowel shift

The eternal circular optimization loop is also seen in the vowel shifts that occurred around 300 years ago in many West-Germanic dialects:



(45)

The arrows denote the direction of the chain shift. Basically, monophthongs rose (Dutch slaːpə ‘sleep’, Old English slæːpɑn, Old Frisian sleːpan, English sliːp), and diphthongs fell (Limburgian wiːn ‘wine’, Dutch ʋɛin, English wain, Flemish waːn). The chain may have started as a result of the lengthening of short vowels in open syllables, which crowded the height dimension if their lengths and tone contours came to equal those of the originally long vowels (Limburgian, which preserves the West-Germanic long vowels, developed a three-way length contrast and a tone contrast).

The rise of the monophthongs can be understood from an asymmetry between adults and children. An adult open mid vowel [ɛː] with an F_1 of 600 Hz will be imitated by young children as a vowel with an F_1 of 600 Hz instead of as an open mid vowel, if we assume that it takes the learner some time to develop an adult-like vowel-height normalization strategy. To implement a vowel with an F_1 of 600 Hz, the child produces a close mid vowel [eː]. If the associated articulation persists into adulthood, the vowel will have risen from one generation to the next. Thus, monphthongs tend to rise if their primary acoustic cue is F_1.

The fall of the first part of the diphthongs can be understood from the contrastive representation of diphthongs. If a language has a single diphthong, its primary perceptual feature may well be its *diphthongal character* (e.g., the presence of an unspecified F_1 fall), by which it is contrasted with all the other vowels. Lowering of the first part of the diphthong amounts to enhancing the contrast with the other vowels: the more the two parts of the diphthong differ from one another, the more they will contribute to the diphthongal character.

6 Conclusion

The strict-ranking approach allows us to model an eternally improving sound system. Even when there are no external factors, sound change may go on forever. The possibility of circular optimization is a property of random constraint variation in general, not of the details of the more involved production grammars that I proposed in §3.

When viewed from a distance, the procedure seems teleological, because the inventory has improved. The changes were, however, automatic, not goal-oriented; the apparent teleology arises because the constraints themselves are functional principles.

References

Andersen, Henning (1969): "Lenition in Common Slavic", *Language* **45**: 553-574.

Archangeli, Diana (1988): "Aspects of underspecification theory", *Phonology* **5**: 183-207.

Beekes, Robert S.P. (1990): *Vergelijkende taalwetenschap*. Het Spectrum, Utrecht.

Boersma, Paul (1989): "Modelling the distribution of consonant inventories by taking a functionalist approach to sound change", *Proceedings of the Institute for Phonetic Sciences of the University of Amsterdam* **13**: 107-123. [available from my home page: http://fonsg3.let.uva.nl/paul/]

Boersma, Paul (1990): "Modelling the distribution of consonant inventories", *Proceedings Linguistics & Phonetics '90*, Prague. [or http://fonsg3.let.uva.nl/paul/]

Boersma, Paul (1997a): "The elements of Functional Phonology", ms. University of Amsterdam. [Rutgers Optimality Archive #173, http://ruccs.rutgers.edu/roa.html]

Boersma, Paul (1997b): "Inventories in Functional Phonology", ms. [Rutgers Optimality Archive #232]

Boersma, Paul (to appear a): "Learning a grammar in Functional Phonology", in: Joost Dekkers, Frank van der Leeuw, & Jeroen van de Weijer (eds.): *The Pointing Finger*.

Boersma, Paul (to appear b): "The OCP in Functional Phonology", ms. [ROA]

Boersma, Paul (to appear c): "Spreading in Functional Phonology", ms. [ROA]

Darwin, Charles (1859): *On the Origin of Species by means of Natural Selection, or The Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Frisch, Stefan, Michael Broe & Janet Pierrehumbert (1997): "Similarity and phonotactics in Arabic", Rutgers Optimality Archive #223.

Dols, Willy (1944): *Sittardse diftongering*. [Alberts' Drukkerijen, Sittard 1953]

Flemming, Edward (1995): *Auditory Representations in Phonology*. Doctoral thesis, University of California, Los Angeles.

Gamkrelidze, T. & V. Ivanov (1973): "Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse", *Phonetica* **27**: 150-156.

Jakobson, R., E.C. Cherry & M. Halle (1953): "Toward the logical description of languages in their phonemic aspect", *Language* **29**: 34-46.

Kawasaki, Haruko (1982): *An Acoustical Basis for Universal Constraints on Sound Sequences*. Doctoral thesis, University of California, Berkeley.

Labov, William (1994): *Principles of Linguistic Change*. Volume I: Internal Factors. Blackwell, Oxford.

Lass, Roger (1994): *Old English: A Historical Linguistic Companion*. Cambridge University Press.

Maddieson, Ian (1984): *Patterns of Sounds*. Cambridge University Press.

Martinet, André (1955): *Economie des changements phonétiques*. A. Francke, Bern.

Moscati, Sabatino, Spitaler, A., Ullendorff, E., & Soden, W. von (1964): *An Introduction to the Comparative Grammar of the Semitic Languages*. Otto Harrassowitz, Wiesbaden. [2nd printing, 1969]

Ohala, John J. & Carol J. Riordan (1979): "Passive vocal tract enlargement during voiced stops", in: J.J. Wolf & D.H. Klatt (eds.): *Speech Communication Papers*, Acoustical Society of America, New York, pp. 89-92. [*JASA* **65 (S1)**: S23]

Ohala, John J. (1993): "The phonetics of sound change", in: C. Jones (ed.): *Historical Linguistics: Problems and Perspectives*, Longman, London, pp. 237-278.

Passy, Paul (1890): *Etude sur les changements phonétiques et leurs caractères généraux*. Librairie Firmin - Didot, Paris.

Prince, Alan & Paul Smolensky (1993): *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report **2**.

Schouten, M.E.H. (1981): "Het verschil tussen *bot* en *bod* — een vergeefse speurtocht", *Nieuwe Taalgids* **74-6**: 537-546.

Sihler, Andrew L. (1995): *New Comparative Grammar of Greek and Latin*. Oxford University Press, New York.

Steenwijk, Han (1992): *The Slovene Dialect of Resia. San Giorgio*. Doctoral thesis, University of Amsterdam.

Steriade, Donca (1987): "Redundant values", in: A. Bosch, B. Need & E. Schiller (eds.): *Papers from the Parasession on Autosegmental and Metrical Phonology*, Chicago Linguistic Society, pp. 339-362.

Streitberg, Wilhelm (1896): *Urgermanische Grammatik*. Carl Winter, Heidelberg.

Stuart-Smith, Jane (1995): "The role of phonetics in the evaluation of reconstructed sound change", *XIIIth International Congress of the Phonetic Sciences* **4**: 682-685.

Trask, R.L. (1996): *A Dictionary of Phonetics and Phonology*. Routledge, London.