[This paper appeared in *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* **13**: 107-123 (1989). For the web version, I modernized the phonetic font and corrected one English word.]

MODELLING THE DISTRIBUTION OF CONSONANT INVENTORIES BY TAKING A FUNCTIONALIST APPROACH TO SOUND CHANGE

Paul Boersma

1. INTRODUCTION

Certain kinds of sound inventories are more widely found in the world's languages than other kinds. Identifying a language with its speakers, we can state that languages show a preference for certain kinds of sound systems at the cost of other ones. Every language has attained its sound inventory in the course of a long-termed development involving many sound changes. In this paper we will adopt the proposition that it is these very sound changes that have let the language acquire the preferable collection of sounds. We will therefore model sound changes as optimizations of sound systems.

2. PRINCIPLES THAT OPTIMIZE SOUND SYSTEMS

According to Passy (1890), sound changes have the same cause that motivates the existence of language itself: people speak in order to be understood, and only to be understood. This functionalist approach manifests itself in two principles:

1) Languages tend to get rid of anything that is superfluous. This principle could be called the *law of least effort*, but Passy prefers (after Sweet) the term *principle of economy*, as it is also assumed to trigger processes like the loss of unaccented vowels, which presumably trades the articulatory effort of the resulting consonant clusters for the smaller time needed to finish the utterance. Among the processes ascribed to this principle are the weakening of accent and the subsequent loss or paradigmatic merger of sounds in unimportant syllables, the simplification of consonant clusters, assimilation, and the abridgement of long vowels.

2) Languages tend to stress or exaggerate anything that is necessary. This is the *principle of emphasis*. First, processes like aspiration and affrication of plosives are considered to be due to this principle, as well as vowel epenthesis, glide insertion, dissimilation, and the change of approximants into fricatives or plosives. The second action of this principle is that of increasing the distinction between two sounds in order to reduce confusion between different words.

   Kawasaki (1982) draws our attention to the acoustic correlates of these two aspects of the principle of emphasis, pointing out that sequences of acoustically similar sounds such as [wu] or [ji] are avoided in the world's languages in favour of sequences with a greater acoustical dynamic variation like [wi] or [ju], and that poorly distinguishable sequences such as [gla] and [dla] tend not to co-occur in languages. In her words, languages tend to maximize *salience* and *dissimilarity*, respectively.

We therefore assume that the development of any sound system aims at three goals:
1) Maximizing the ease of articulation.
2) Minimizing confusion in the vocabulary by maximizing the perceptual distinctions between words.
3) Maximizing the perceptual salience within words, i.e. maximizing the perceptual contrast between adjacent sounds in the speech chain.

3. THE WAYS OF SOUND CHANGE

A traditional view of sound change regards it as a gradual process (e.g. Hockett, 1965). The reason for this is that speakers seem to be unaware of ongoing sound changes. But then, speakers are unaware of quite a lot of other parts of their grammar. And so, many writers on this subject reject the graduality hypothesis (e.g. Hoenigswald, 1964). According to Passy (1890), who lays the source of sound change in children imperfectly trying to master the language that surrounds them, the function of language is not impaired as long as the mispronouncing child's utterances are still being understood and are not too strongly disapproved of by her environment. Chomsky and Halle (1968, p. 249-252), who place the origin of language change in the adult speech community, view sound change as the addition of a rule to the grammar, which is then restructured by the next generation of children learning the language.

Identification of the speakers that change their speech will prove not to be crucial to the present paper, but taking up a position on the graduality question is required. We will assume with Passy that *natura facit saltus*.

4. MEASUREMENT OF OPTIMIZATION

The degree to which a certain sound change constitutes an optimization in articulatory effort, is a function of the concerning tracks through articulatory space. Likewise, an evaluation of perceptual salience measures a function of the path that the speech utterance takes through perceptual space. Finally, the dissimilarity between two utterances is a function of the respective tracks through perceptual space. It will therefore be rewarding to discriminate between articulatory and perceptual space. This is done in the next two sections with special attention to obstruent consonants.

4.1. The articulatory space

We have the following degrees of closure, classified according to perceptual differences, i.e., every pair of successive non-parenthesized labels is found somewhere in the world to contrast two phonemes:

Labial, coronal and dorsal opening (the areas given have been extracted from the figures in Fant (1960)):

0	Complete closure, as in stops consonants.
(1)	Stop release burst.
2	Incomplete closures, as in fricatives.
3	Approximants and strong secondary articulation.
4	Vocalic constrictions: high vowels, glides, liquids, lip rounding; 0.3-1 cm^2.
5	Neutral vocalic opening (mid vowels); 1-4 cm^2.
6	Large vocalic opening (spread lips, low vowels); 4-15 cm^2.

Nasal opening (numbers having the same areas as above):
0 Complete closure, no nasality.
5 Open nasal tract; 2 cm^2 in Fant (1960).

Pharyngeal opening (numbers having the same areas as above):
2 Pharyngeal fricatives.
3 Heavily pharyngealized vowels and consonants.
4 Vocalic pharyngeal constriction, as in [a]; 0.3-1 cm^2.
5 Neutral vocalic opening, as in the 'lax' dorsal vowels [ɛ], [ɪ], [ɔ] and [ʊ].
6 Advanced tongue root, as in the 'tense' dorsal vowels [e], [i], [o] and [u].

Glottal opening:
0 Constricted glottis, as in the glottal stop and ejectives.
1 Narrowed glottis; a necessary condition for voicing. Also present in voiceless unaspirated plosives (Slis & Damsté, 1967).
2 Slightly spread glottis, as in [ɦ] and breathy voiced consonants.
3 Spread glottis, as in [h] and aspirated consonants.

We have the following degrees of tension:

Supralaryngeal tension (mainly the pharyngeal constrictor muscle, but the oral cavity walls play a part as well):
0 Lax.
1 Tense.

Vocal cord tension:
0 Slack; a necessary condition for plosive voicing. In sonorants this causes a low pitch.
1 Neutral.
2 Stiff vocal cords. Causes a high pitch in sonorants.

Table 1 shows the articulatory features of a number of sounds that involve a constrictive gesture of the lower lip. A characteristic degree of underspecification is shown where the minimum and maximum opening are given. In order that perceptual invariance be maintained, constrictions in front of the main constriction must be appreciably wider than the main constriction, and constrictions behind it must be as least as wide; furthermore, underspecified constrictions must not be narrower than the approximant variety, in order not to introduce the extra feature of friction noise. These considerations account for the

Table 1. Articulatory specification of some labial sounds.
'3-6' means '3 or 4 or 5 or 6'. '|' denotes a temporal contour.

	p	f	v	b	m	w	p^h	v̞	w̃	b̥	b̤	v̥	v^+	ɸ	h^w	u
labial opening	0	2	2	0	0	4	0|3-6	3	4	0	0|3-6	2	2	3	4	4
coronal opening	3-6	3-6	3-6	3-6	3-6	4-6	3-6	3-6	4-6	3-6	3-6	3-6	3-6	3-6	4-6	5-6
dorsal opening	3-6	3-6	3-6	3-6	3-6	4-6	3-6	3-6	4-6	3-6	3-6	3-6	3-6	3-6	4-6	4
nasal opening	0	0	0	0	4	0	0	0	4	0	0	0	0	0	0	0
pharyngeal opening	3-6	3-6	3-6	3-6	3-6	4-6	3-6	3-6	4-6	3-6	3-6	3-6	3-6	3-6	4-6	6
glottal opening	1-2	2-3	1	1	1	1	3	1	1	1	2	2-3	1	2-3	2-3	1
supralar. tension	1	1	0	0	0-1	0-1	1	0	0-1	0	0	0	1	1	1	0-1
vocal cord tension	1-2	1-2	0-1	0	0-2	0-2	0-2	0-2	0-2	1-2	0	1-2	0-1	1-2	1-2	0-2

coronal, dorsal and pharyngeal openings in table 1. In this view, vowels should be more fully specified for supralaryngeal closure features than consonants. Of course, the degree of underspecification of each sound depends on the grammar of the language in question, which has to respect safety margins between opposing phonemes. The form of underspecification presented here is meant to allow for combinatorial variation in languages without phonologized secondary articulations on the one hand or underspecification of manner and place on the other.

4.2. Articulatory effort

Articulatory effort is a function of the sounds in an utterance, measured in the articulatory space. Table 2 shows possible articulatory gestures for the sound sequences [ampa] and [amta]. The difference between these two utterances as to ease of articulation lies in the [mp] and [mt] transitions. In [mp] the only gesture is the closing of the nasal tract, whereas in [mt] there are three gestures that have to coincide more or less. As a strict simultaneity is impossible, this can only be achieved by choosing one out of six ways in which these three contours can be ordered. This choice depends on the grammar of the language in question, especially on syllabification rules and on how much importance the language attaches to several kinds of perceptual invariance.

Table 2. The articulation of [ampa] and [amta].

	a	m	p	a	a	m	t	a
labial opening	6	0	0|1	6	6	0	3|6	6
coronal opening	6	6	6	6	6	6|3	0|1	6
dorsal opening	6	6	6	6	6	6	6	6
nasal opening	0	5	0	0	0	5	0	0
pharyngeal opening	4	4	4	4	4	4	4	4
glottal opening	1	1	1	1	1	1	1	1

4.3. The perceptual space

As a raw discretization of the perceptual space, we stipulate the existence of the binary perceptual features voiced, fricative, continuant, nasal, tense, vocalic. Place features will be neglected for the moment. The perceptual features can be obtained from the articulatory features as follows:

- *Supralaryngeal opening*: this is the conductance of the vocal tract above the larynx and can be approximated as the minimum of the pharyngeal opening and the suprapharyngeal opening, which in turn can be approximated as the maximum of the oral opening and the nasal opening. The oral opening, again, is the minimum of the labial, the coronal and the dorsal opening.

- *Voiced*: the vocal folds will only vibrate if the situation satisfies something like

  laryngeal pressure drop > (glottal opening) * (vocal cord tension) * (a constant)

  in real numbers. In our notation, this condition is fulfilled if the glottal opening is labelled '2' and the vocal cords are held slack, or if a glottal opening of '1' coincides with either of the following conditions:

- the supralaryngeal opening is at least '2'.
- this opening is smaller than '2' and the vocal cords are held slack.

- *Tense*: in tense voiced plosives, voicing will cease earlier than in lax voiced plosives; tense unvoiced plosives will have a stronger release burst than lax ones. Tenseness, therefore, is a perceptual feature for plosives.

- *Continuant*: this feature is present only if there is an oral opening throughout.

- *Fricative*: this feature relates to prominent friction noise; it is present:
  - if the supralaryngeal opening is '2' and voicing is present: [v], [β].
  - if this opening is at least '2' and there is no voicing: [f], [ɸ], [hʷ], as well as the second part of [pʰ].
  - with breathy voicing: [b̤].

- *Sonorant*: this feature is present only if the supralaryngeal opening is at least '4'.

- *Nasal*: this feature is present only if the nasal opening is not '0'.

Table 3 shows the perceptual features of all the sounds that appeared with their articulatory features in table 1.

Table 3. Perceptual features of some labial sounds.

	p	f	v	b	m	w	pʰ	v̞	w̃	b̥	b̤	v̥	v⁺	ɸ	hʷ	u
voiced	–	–	+	+	+	+	–	+	+	–	+	–	+	–	–	+
continuant	–	+	+	–	–	+	–	+	+	–	–	+	+	+	+	+
nasal	–	–	–	–	+	–	–	–	+	–	–	–	–	–	–	–
fricative	–	+	+	–	–	–	+	–	–	–	+	+	+	+	+	–
sonorant	–	–	–	–	+	+	–	–	+	–	–	–	–	–	–	+
tense	+	+	–	–	–	–	+	–	–	–	–	–	+	+	+	–

The sounds are also shown in figure 1, with lines connecting sounds that differ in only one perceptual feature. Sound changes are supposed to run along these lines, but larger jumps are probably possible.

Figure 2 (next page) shows the 'cardinal' labial consonants; the lines connect sounds that differ in exactly *two* perceptual features; the consonants that are not connected by a line differ in four features. One of the things to be learnt from this figure is the fact that a

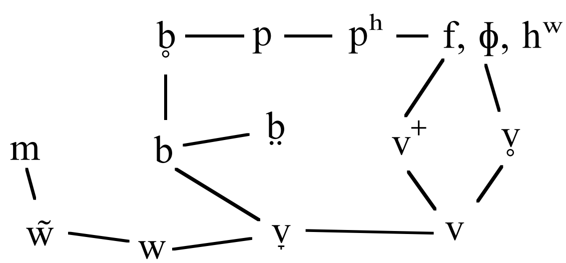

Figure 1. Labial sounds. Lines connect sounds that differ in exactly one feature.

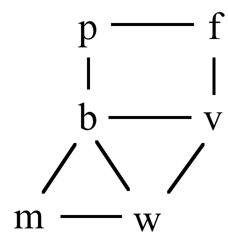

Figure 2. The 'cardinal' labial consonants. Lines connect sounds that differ in exactly two features.

glide and a voiced stop are perceptually neighbouring sounds that have in common the absence of friction noise.

4.4. Perceptual salience

In first approximation, the perceptual salience of a sound sequence can be measured as the number of perceptual features changing their values in lapse of time. There is no difference in features between [w] and [u], and a difference of five features between [p^h] and [u], so the salience of [wu] is very small and that of [p^hu] quite large. The fact that maximizing salience is a productive motive in sound change, is shown by the instance of the Low Franconian dialects of the Belgian province of Limburg. When these languages raised [ɔː] into [oː] (independent of tone), the word [wɔː^2] was apt to lose much of its internal salience but was rendered [boː^2] in some regions and [moː^2] in others in order to maintain or restore the original salience, whereas [w] did not change in other positions (Tans, 1938; the sign ^{2} denotes falling tone). The same reasoning explains why Latin chose to have *bōs* 'cow' instead of the regular *[woːs].

Place features play a role in determining the salience of a sound sequence, as can be seen in table 4.

Table 4. Perceptual place contours in plosive + vowel sequences

	pa	pi	pu	ta	ti	tu	ka	ki	ku
labial	+ -	+ -	+	–	–	+	–	–	+
coronal	–	–	–	+ -	+ -	+ -	–	–	–
dorsal	–	+	+	–	+	+	+ -	+	+

It appears that [t] enforces a coronal contour with all three vowels [a], [i] and [u], whereas [p] yields only two contours, and [k] no more than one. This correlates with the fact that the coronal articulator is more frequently used in the world's languages than the other articulators: Maddieson (1984) counts 27 languages that lack [t] (dental or alveolar), 34 languages without [k] and 54 languages without [p], in a sample of 317 languages; the flexibility and rapidity of the apical articulator probably play a role as well.

4.5. Perceptual dissimilarity

As a first approximation, the difference between two sounds is measured by the number of features that are distinct. This approximation may help us in deciding the possible paths of sound change, but it seems much too crude to measure the perceptual confusion between words. For our purposes it is necessary that we can compare dissimilarities of different pairs in order to work out whether or not a proposed sound change improves system distinctivity. Data from experiments on confusion probabilities are liable to be of use here, though care must be taken of their language dependencies. These things are outside the scope of the present paper, however.

4.6. Favoured sounds

Words that will be favoured by the world's languages without respect to the rest of the vocabulary, are words that require little articulatory effort and feature much perceptual salience. Consider, for instance, the sound sequence [ata].

To pronounce this sequence, the speaker makes a coronal closing and opening gesture only and does not have to adjust any other closures, nor the active state of her glottis. She might find it convenient to make a light concurrent movement with her mandible, but that could be all. In articulatory space, therefore, the [t] in [ata] lies very close to [a] and articulatory effort for this sequence is low.

The situation is quite different in perceptual space, however. The primary constrictions of [t] and [a] are at quite distinct places, and these sounds differ in no less than four manner features. The salience of the [ta] sequence is accordingly quite large. And so, from the viewpoint of maximal ease and salience it is strongly favoured that consonants and vowels should alternate in the speech chain.

For instance, from the three consonants [p], [t] and [k] and the three vowels [a], [i], [u] we have 27 words on a CVCV-pattern. Allowing sequences with different vowels like [ita] greatly increases the information that can be conveyed in a certain lapse of time (81 possible words). As the articulatory changes from [i] to [a] can be enforced during the stop closure interval and pose no strong requirements on simultaneity, allowing such sequences could be the next step in creating a phonology. That this possibility has not been exploited in *all* grammars can be seen from many CV-syllabled languages disallowing or disfavouring within a word, changes in vowel backness, rounding and/or *advanced tongue root*.

5. THE INTERACTION BETWEEN THE OPTIMIZATION PRINCIPLES

How does a language decide if a sound change, which proposes itself by always being present in the variations inside the speech community, should be granted access to the grammar? The first strategy that comes to mind is:

1) Compute the values of the current and the proposed sound system as a function of the above three principles:

    Value = Ease + Salience + Distinctivity

2) Allow the proposed change to take place if the value of the proposed sound system is greater than the value of the current sound system.

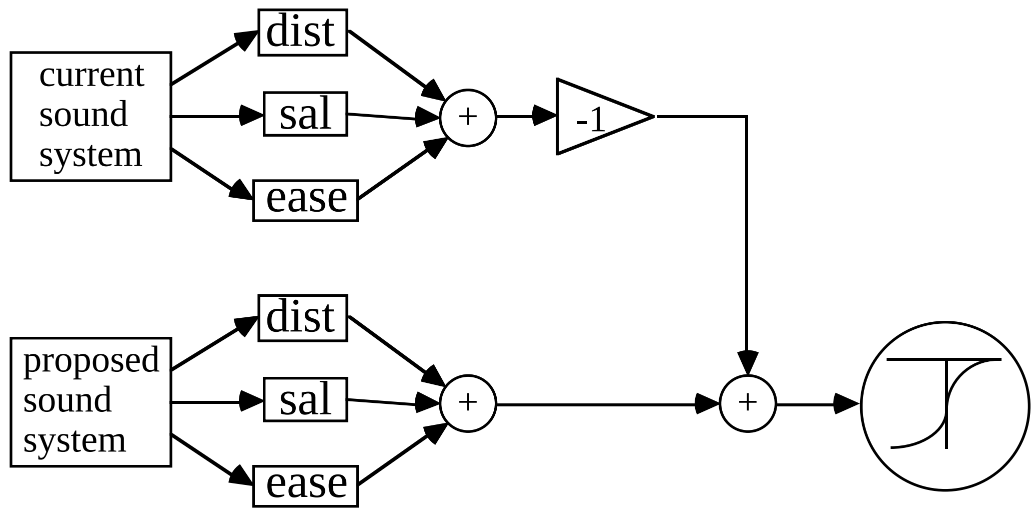

Figure 3. Decision flow for the bookkeeper's strategy for allowing a sound change.

This strategy is depicted in figure 3. It requires that numbers be assigned to the ease, salience and distinctivity in a sound system and that the relative importances of these three principles are established.

A strategy that does not presume nearly as many parameters and calculations, contains the following steps (see figure 4):

1) See if the distinctivity of the proposed sound system is better than that of the current sound system. In the case of an unconditioned sound change, this boils down to looking for the nearest neighbours of the changing sound before and after the change and comparing the distances or confusion probabilities.
2) See if the salience improves.
3) See if articulatory ease improves.
4) Allow the sound change to take place if at least two of the three questions posed above have been answered positively.

In this strategy, which is presented in figure 4, it is possible not to refer to any data measured in numbers. Instead, we can do with a number of rank orderings, as will be seen in our example, where we adopt this very strategy. It is obviously not as good as the first strategy at finding the optimal system, but then, no language seems to be very good at it, as is demonstrated by their refusal ever to stop changing.

6. RELATION BETWEEN INVENTORY DISTRIBUTION AND SOUND CHANGE

The sound inventory **I** of a language is the collection **F** of its phonemes, together with their number **N**. The a priori probability that a certain language possesses the inventory $\{\phi_1, ..., \phi_n\}$, where $\phi_1...\phi_n$ are different elements of the set Φ of all possible speech sounds, is given by

$$\mathbf{P}\big[\mathbf{I}=\{\phi_1,...,\phi_n\}\big] = \mathbf{P}[\mathbf{N}=n] \cdot \mathbf{P}\big[\mathbf{F}=\{\phi_1, ...,\phi_n\} \mid \mathbf{N}=n\big] \qquad (1)$$

The probability that next year's inventory **I'** will have the value J is

$$\mathbf{P}[\mathbf{I}'=J] = \sum_I \mathbf{P}[\mathbf{I}'=J \mid \mathbf{I}=I]\, \mathbf{P}[\mathbf{I}=I] \qquad (2)$$

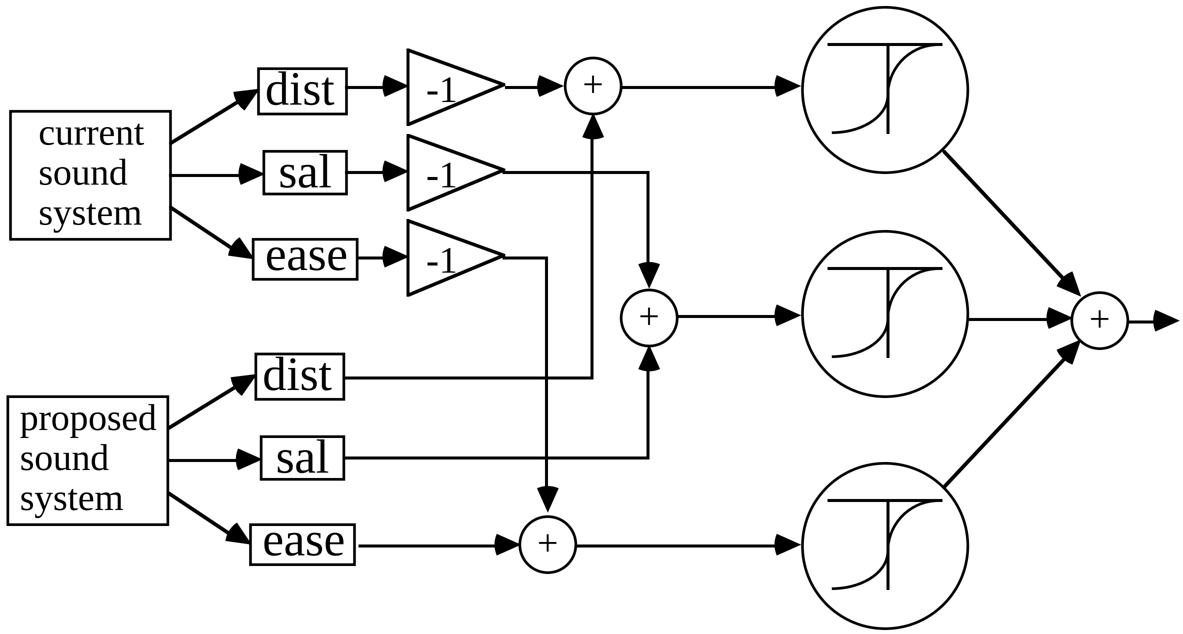

Figure 4. A 'democratic' model of decision flow.

where **I** is this year's inventory and **P**[**I**'=*J* | **I**=*I*] denotes the probability that an inventory *I* will change into an inventory *J* in the course of a year. Or, if we define the probability vector

$$p_I \equiv \mathbf{P}[\mathbf{I}=I] \tag{3}$$

and the transition matrix

$$T_{JI} \equiv \mathbf{P}[\mathbf{I}'=J \mid \mathbf{I}=I] \tag{4}$$

then we can rewrite eq. (2) as

$$p_J' = \sum_I T_{JI}\, p_I \qquad \text{or} \qquad p' = T\, p \tag{5}$$

An equilibrium will be obtained when languages will have grown very old. In that case

$$p = T\, p \tag{6}$$

Therefore, the equilibrium distribution is an eigenvector of the transition matrix. Its eigenvalue equals unity. The dimension of the pertinent eigenspace equals the number of independent partitions of the collection of all possible speech sounds. If, for instance, the distribution of vowels in languages is independent of the distribution of consonants, the dimension of the space of positive-definite eigenvectors of the transition matrix is at least two. The most general transition matrix has thousands of rows and columns, but we shall presently have a chance to calculate the equilibrium distribution of a set of inventories that consist of phonemes that can be approximated as being nearly independent in distribution from other phonemes.

7. AN EXAMPLE: THE GERMANIC CONSONANT SHIFT

The most published upon of all known sound changes is the *Germanic consonant shift* (data in Brugmann & Delbrück, 1897; a different analysis in Meillet, 1922). Schmitt (1949) gives an account that attributes all of the changes [p]→[ɸ]→[β]→[b]→[p] to the law of least effort (all these changes are attested: for the first change, compare Latin *pellis*, *tre:s*, *canis* 'skin, three, dog' to Germanic *fell-*, *θri-*, *hund-* 'skin, three, dog'; for the third, compare Latin *fru:ctus* < *β-*, *facio* < *δ-*, *hortus* < *γ-* 'enjoyment, do, yard' to Germanic *bru:k-*, *do:-*, *gard-* 'use, do, yard'; for the fourth, compare Latin *duo*, *gelu* 'two, frost' to Germanic *two:*, *kald-* 'two, cold'; these three changes are called Grimm's law; the second change is referred to as Verner's law and is responsible for accent-dependent voiced-unvoiced alternations in early Germanic fricatives). His argument runs as follows: a [ɸ] is easier to pronounce than a [p], because the lips do not have to travel as far. A [β] is easier to pronounce than a [ɸ], because in pronouncing a [β] it is no longer the cheeks, but the glottis that resists the pulmonary pressure. A [b] is easier than a [β], because the articulation of a stop consonant is allowed to be less precise than the articulation of a fricative. This argument has been supplied with several metaphors in the literature. The reverse argument could have been invoked *against* [p]→[ɸ], but [ɸ] needs less precision of articulation than [β] does, as a labial approximant will sound as a fricative if it is voiceless, but not if it is voiced. Finally, [p] is pronounced with less effort than is [b], because retaining voicing during complete supralaryngeal closure must invoke the additional articulatory gesture of lowering vocal fold tension in order to compensate for the decreased excess pulmonary pressure.

Schmitt's approach is summarized in table 4. Each of the four *rows* shows one improvement and two deteriorations. Thus, if the four factors that constitute articulatory effort are assigned equal importance, the improvement in each *column* is greater than any of the deteriorations in the same column. The decision strategy that Schmitt tacitly assumes, therefore, is yielding to the loudest cry, as opposed to ours, which is counting votes.

Table 4. Optimization of articulatory effort in the Germanic consonant shift according to Schmitt (1949).

	[p] → [ɸ]	[ɸ] → [β]	[β] → [b]	[b] → [p]
precision:	worse	worse	better	–
excursion:	better	worse	worse	–
voicing:	worse	–	worse	better
tenseness:	worse	better	–	worse

We shall now see how the three optimization principles of articulatory ease and perceptual salience and distinctivity co-operate on obstruent consonants.

As an example of a subset of the collection of all possible consonant inventories, we consider the set of inventories that consist of three labial obstruents out of {p, b, f, v, ph}. This collection of sounds is chosen on the ground that it contains a large number of obstruents (five) under the condition that every pair of them can co-occur as different phonemes in the same language.

The set of possible consonant inventories now contains ten different inventories. We further take {a} as our vowel inventory and make stressed CV-words accordingly. All languages that are possible only contain three of the five possible words [pa], [ba], [fa], [va], and [pha].

It is supposed that sounds change one at a time along the lines p-ph-f-v-b-p either way.

Optimization is measured as follows.

1) Articulatory effort:
- 1a) Vowels are spontaneously voiced: [pa] > [pha]; [fa] > [pha]. The sign '>' means here 'is better than' in the respect of articulatory ease.
- 1b) Plosives are easier than fricatives: [ba] > [va].
- 1c) Plosives are best not voiced: [pa] > [ba].
- 1d) Non-plosives like to be voiced: [va] > [fa].

Schmitt's reasoning is used in 1b, 1c and 1d. Statements 1a and 1b are ordered with respect to each other. In a certain sense, 1b is ordered with respect to 1c and 1d as well.

2) Perceptual salience:

If we count the number of perceptual features that distinguish the consonant from [a] and take this number as a measure of perceptual salience, we find the ordering [pha] > [pa, fa] > [ba, va] . Before a stressed vowel, continuancy is probably a perceptually more important feature than friction, so [ba] > [va] . Incidentally, the order that we find is reflected in the sonority hierarchy.

3) Perceptual dissimilarity:

We assume that:
- 3a) The dissimilarity of any pair of non-adjacent sounds from {p, b, f, v, ph} is greater than the dissimilarity of any pair of adjacent sounds.
- 3b) The a priori voicing distinction is larger for labial plosives than it is for labial fricatives.
- 3c) The plosive-fricative distinction is larger for voiceless obstruents than it is for voiced obstruents. This statement is contained in the combination of 3a and our choice of three unvoiced obstruents opposing two voiced ones.
- 3d) The distinction between [b] and [v] is larger than that between [ph] and [f].
- 3e) The distinction between [f] and [v] is larger than that between [p] and [ph].

These results are recapitulated in figure 5.

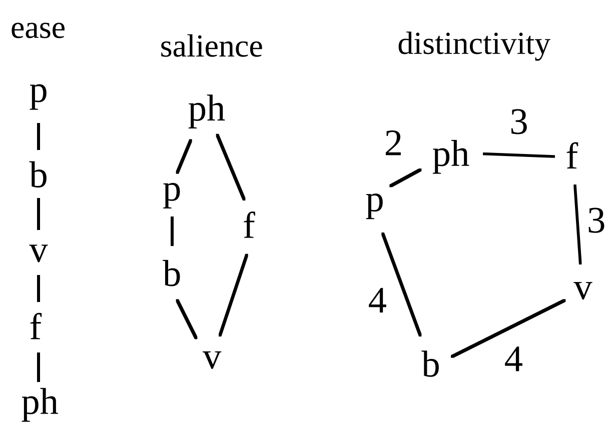

Figure 5. Ordering relations for labial obstruents before an accented vowel. For the distinctivity, a possible two-dimensional scaling is shown. In a Euclidean metric, all distinctivities of non-adjacent pairs are greater than 4.

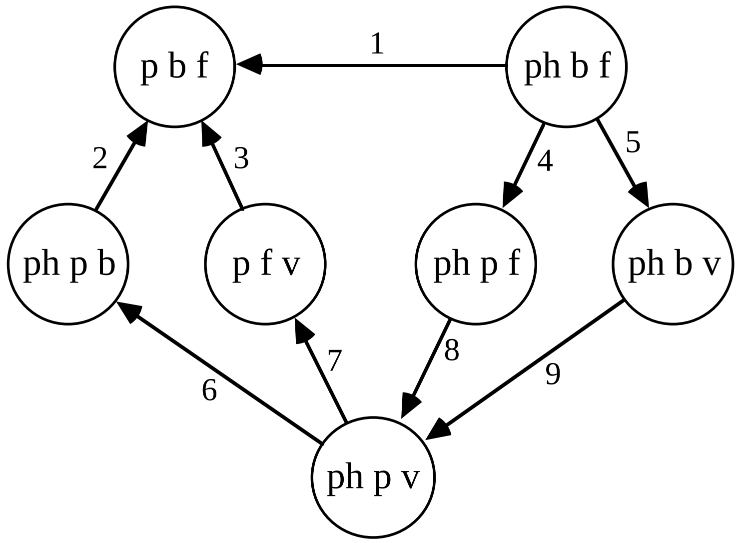

Figure 6. The possible changes inside the set of inventories of three labial obstruents in 'strong' position (before a stressed vowel). All these sound systems will eventually convert into a {p,b,f}-system. In order to maintain an analogy with figure 9, the inventories {ph,f,v}, {p,b,v} and {b,f,v} are not shown.

This works out for sound inventories in the following way. A {*pa*, *ba*, *va*} language, for instance, will become a {*pa*, *ba*, *fa*} language because the changes in salience and ease cancel each other, but the dissimilarity of [fa] and [pa] is greater than that of [va] and [ba]. Or, alternatively, it may become a {*pha*, *ba*, *va*} language as this change would involve an increase in salience and distinctivity. As these two processes exhaust the possibilities of {*pa*, *ba*, *va*}, this sound inventory will vanish from the languages of the 'world'. A {*pha*, *ba*, *va*} language, however, can transgress into {*pha*, *pa*, *va*} as that increases salience and ease, and the latter one can in its turn be converted into {*pha*, *pa*, *ba*}. As figure 6 shows, the {*pa*, *ba*, *fa*} inventory eventually emerges as the sole survivor, and that will be the end of sound change.

It would be the end of our story, too, were it not for an asymmetry in the speech organs that comes to rescue. The only reason, after all, that a {*pha*, *ba*, *fa*} language will become a {*pa*, *ba*, *fa*} language, is that [pa] is more unlike [ba] than [pha] is as compared to [fa]. The analogous relation might not hold for other articulators. So let us now take a look at figure 7.

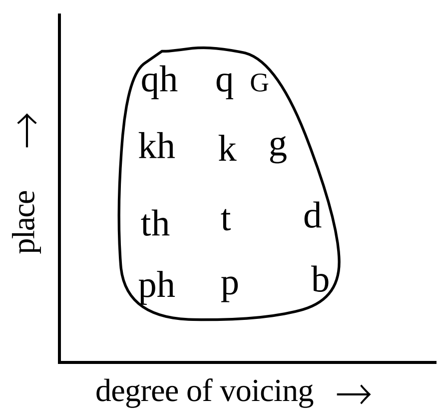

Figure 7. A two-dimensional projection of the perceptual space of plosives.

It shows a projection of the perceptual space of plosive consonants onto axes that represent information on place and on voicing. Voicing can be retained longer in [b] and [d] than it can in [g], because the volume of the cavities behind the constriction is smallest in [g]. This causes a faster decrease in the glottal pressure drop and consequently an earlier ceasing of vocal cord vibrations in [g]. The data of Bothorel-Witz (1978) show this effect for Stage German. There are quite a lot of languages that have [t], [d] and [k] but lack [g]. In fact, in Maddieson (1984) we have the following numbers of coronal and velar plosive systems ('*' denotes a gap):

88	t	k
	d	g

4	t	k
	*	g

10	t	k
	d	*

	th	kh
0	t	k
	*	g

	th	kh
7	t	k
	d	*

	t’	k’
0	t	k
	*	g

	t’	k’
6	t	k
	d	*

The large probability of confusing [k] and [g] has orchestrated several sound changes:

- In Japanese (McCawley, 1968), there was no velar nasal, so /g/ developed a sonorant allophone [ŋ] intervocalically, which remains an underlying obstruent synchronically.

p	t	k
b	d	g
		↓
m	n	

- In Low Franconian (Van Loey, 1921), there existed voiced fricatives, but no [ɣ]. This was the only hole, as a velar nasal was being derived from the sequence /Ng/. The /g/ developed a fricative allophone, first intervocalically and eventually in every position. Only a few dialects now retain [g] in the former geminate position.

p	t	k
b	d	g
		↓
v	z	
m	n	(ŋ)

It is not uncommon that [g] goes its own way:

- The Proto-Semitic /g/ developed in Arabic into a palato–alveolar affricate in every position, whereas /k/ did not change (Moscati et al., 1964):

	t	k
b	d	g

→

	t	k
b	d	dʒ

Or it can drag [b] and [d] along with it:

- In Castillian (Llorach, 1950; Harris, 1969), all voiced plosives developed into non-strident fricatives in most positions:

p	t	k
b	d	g

→

p	t	k
β	ð	ɣ

So what happens in the model proposed in section 5? A chart for velar obstruents that is analogous to figure 5 is shown in figure 8. The similarity between [k] and [g] is greater than that between [p] and [b], but that is the only difference between figures 5 and 8. The flow of velar obstruent systems is shown in figure 9: if [kha] is more different from [xa] than [ka] is from [ga], a {*ka*, *ga*, *xa*} language will change its [ka] to [kha]. With the assumptions laid down in figure 8 and our three-way decision strategy, there will be no optimal velar obstruent system.

If in a certain language the velar obstruents follow the paths depicted in figure 9 and the labial and coronal obstruents follow by analogy (or due to the ease inherent in articulatory organization), this language's solution to the [g]-problem is

p	t	k
b	d	g

→

ph	th	kh
b	d	g

Subsequently, [b, d, g] will be devoiced to yield [p, t, k], because this change improves salience and articulatory effort, and so on.

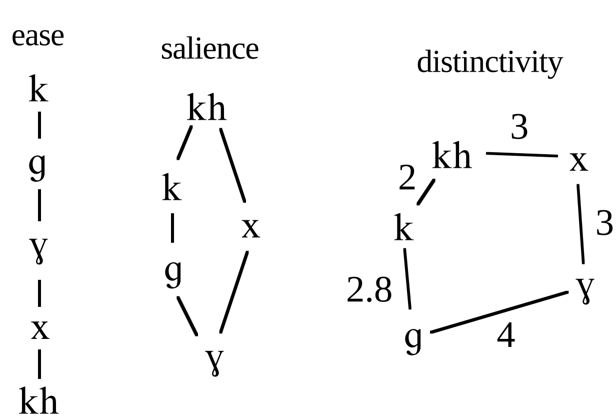

Figure 8. Optimization orderings and a dissimilarity scaling for velar obstruents.

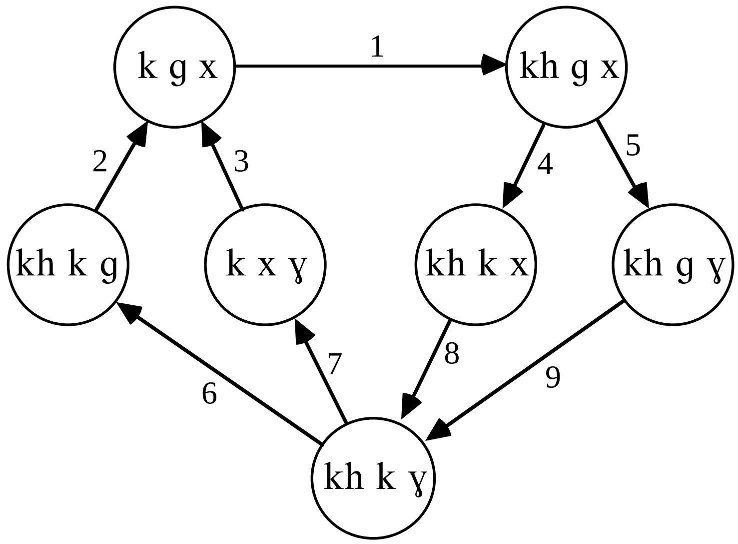

Figure 9. Circular optimization for inventories containing three velar obstruents.

A count of the sound systems in Maddieson supports our claim that the presence of a [g] is needed for the voiceless plosives to become aspirated, though the results are not very significant ($\chi^2 = 1.5$, $p = 0.11$):

88	t	k	10	t	k	13	th	kh	0	th	kh
	d	g		d	*		d	g		d	*

The changes depicted in figure 9 are the Germanic consonant shift again. The original {β, p, b} system became {f, p, b} in Old Latin, but changed to {β, ph, b}, {β, ph, p} and finally {β, f, p} in Proto-Germanic (the order was not recorded), both consistent with our model. In Germanic this {β, f, p} became {b, f, p}, the best system according to figure 6. It is still this in a lot of Low-German dialects that lost [g] along the way. The Germanic dialects that retained [g], however, changed the {b, f, p} system into {b, f, ph} (Swedish), or further into {p, f, ph} (Danish), or still further into {p, v, pf} (many High German dialects). Things have been complicated by changes of place like [x] → [h], changing sonorants like [w] → [v], and mergers as there are [θ, t] → [t] and [ð, d] → [d], but the general idea of the second Germanic consonant shift is reflected in our model.

So we have arrived at an important conclusion. Under the three-decision scenario we have the possibility of a sound inventory that never stops optimizing. The circularity is quite resilient. For instance, if distinctivity condition (3d) is not valid, branch 5 in figure 9 will be reversed; if condition (3e) does not hold, branch 7 will be reversed. Neither case, however, will remove the circularity of the optimization track. Conditions (3b) and (3c) play no role whatsoever. Only condition (3a), i.e. the choice of the sound system, and the condition governing branch 1 are crucial to the circularity. Furthermore, expanding the sound inventories will decrease the likelihood of the existence of an optimal sound system.

If equal probabilities are assigned to all transitions in figure 9, the equilibrium is found to have 25% of the three-obstruent languages in {k, g, x} and 12.5% in each of the inventories {kh, g, x}, {kh, k, x}, {kh, g, ɣ}, {kh, k, ɣ}, {kh, k, g}, and {x, k, ɣ}.

No language is predicted to have any of the systems {kh, x, ɣ}, {x, ɣ, g}, or {k, g, ɣ}. A count of the three-obstruent sytems in Maddieson yields the results of table 5. The {p, b, f}-like systems abound for labials as well as for dorsals, but not for coronals, which is probably due to the abundance of [s] in the world's languages, which strongly frustrates the existence of [θ]. Inventories without any non-sibilant fricative seem fairly well attested. The prominence of the coronals in this respect is not due to an overall preference for them in language, but to the fact that many {t^h, t, d}-languages have more than three labial or dorsal obstruents and/or lack a [g]. The relatively high rate of occurrence of {p, b, v}-like systems with all three articulators could be due to the accent system being different in many languages from the omnipotent stress accent underlying the sound changes in our model; in particular, the perceptual dissimilarity of [p] and [p^h] greatly diminishes in an unstressed environment.

Table 5. The frequency of inventories of three labial, coronal or dorsal obstruents in Maddieson. The inventory was not registered if it additionally contained an implosive or ejectives. [s]-like phonemes were not taken into account.

labials					coronals					dorsals				
17	*	p	b	f	1	*	t	d	θ	13	*	k	g	x
4	*	p^h	b	f	0	*	t^h	d	θ	3	*	k^h	g	x
5	*	p^h	p	f	0	*	t^h	t	θ	4	*	k^h	k	x
1	*	p^h	b	v/β	0	*	t^h	d	ð	0	*	k^h	g	ɣ
0	*	p^h	p	v/β	0	*	t^h	t	ð	2	*	k^h	k	ɣ
2	*	f	p	v/β	1	*	θ	t	ð	7	*	x	k	ɣ
8	*	p^h	p	b	14	*	t^h	t	d	7	*	k^h	k	g
7	*	p	b	v/β	3	*	t	d	ð	3	*	k	g	ɣ
0	*	p^h	f	v/β	0	*	t^h	θ	ð	1	*	k^h	x	ɣ
1	*	b	f	v/β	0	*	d	θ	ð	0	*	g	x	ɣ

8. CONCLUSION

In this paper we developed a three-way decision regime that 'only' requires knowledge of rank orderings of the articulatory ease and the perceptual salience of sound sequences and knowledge of the orderings of dissimilarities of pairs of words. Under this regime the sound patterns of languages will keep changing forever, even if there are no external influences on them. The direction of the Germanic consonant shifts, for example, followed from the orderings and the decision strategy that we used.

REFERENCES

Bothorel - Witz, A. (1978): *La sonorité des consonnes corrélatives d'allemand*, Thèse presentée à l'Université des Sciences Humaines de Strasbourg.

Brugmann, K. & Delbrück, B. (1897): *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen*, vol. I, Einleitung und Lautlehre. Trübner, Strasbourg.

Chomsky, N. & Halle, M. (1968): *The Sound Pattern of English*. Harper and Row, New York.

Fant, G. (1960): *Acoustic Theory of Speech Production*. Mouton, The Hague.
Harris, J.W. (1969): *Spanish Phonology*. MIT Press, Cambridge, Ms.
Hockett, C.F. (1965): Sound Change, *Language* 41-2: 185-204.
Hoenigswald, H.H. (1964): Graduality, Sporadicity, and the Minor Sound Change Processes, *Phonetica* 11: 202-15.
Kawasaki, H. (1982): *An Acoustical Basis for Universal Constraints on Sound Sequences*. Dissertation, University of California, Berkeley.
Loey, A. van (1921): *Schönfeld's historische grammatica van het Nederlands*. Thieme, Zutphen. 5th edition, 1954.
Llorach, E. A. (1950): *Fonología española* Editorial Gredos, Madrid. 3rd edition, 1961.
Maddieson, I. (1984): *Patterns of Sounds*. Cambridge University Press.
Martinet , A. (1955): *Economie des changements phonétiques*. A. Francke, Bern.
McCawley, J.D. (1968): *The Phonological Component of a Grammar of Japanese*. Mouton, The Hague.
Meillet , A. (1922): *Introduction à l'étude comparative des langues indo-européennes*. Hachette, Paris.
Moscati, S.; Spitaler, A.; Ullendorff, E. and Soden, W. von (1964): *An Introduction to the Comparative Grammar of the Semitic Languages*. Otto Harrassowitz, Wiesbaden. 2nd printing, 1969.
Passy, P. (1890): *Etude sur les changements phonétiques et leurs caractères généraux*. Librairie Firmin - Didot, Paris.
Schmitt, A. (1949): Zur germanischen und hochdeutschen Lautverschiebung. *Zeitschrift für Phonetik* 3: 1-25.
Slis & Damsté (1967): Transillumination of the Glottis during Voiced and Voiceless Consonants. *IPO Annual Progress Report* 2: 103-109.
Tans, J.G.H. (1938): *Isoglossen rond Maastricht*. Proefschrift Katholieke Universiteit Nijmegen.