Perceptual similarity and articulatory reinterpretation as a source of phonological innovation

Jonasson, J.

journal: STL-QPSR
volume: 12
number: 1
year: 1971
pages: 030-042

http://www.speech.kth.se/qpsr
D. PERCEPTUAL SIMILARITY AND ARTICULATORY REINTERPRETATION AS A SOURCE OF PHONOLOGICAL INNOVATION

J. Jonasson*

0. Contents.
1. The goals of phonological theory.
2. Articulatory reinterpretation can occur when different articulatory gestures have similar perceptual effects.
3. Some cases where articulatory reinterpretation seems to have played a role.
4. Can marking conventions handle reinterpretation?

1.1 Phonological theory should enable us to give an adequate description of phonological processes. As pointed out by Halle (1964a, p. 329) the choice of a specific set of distinctive features "must be understood as a statement about the nature of human language in general", i.e. it is assumed that a restricted set of distinctive features is capable of defining the segments of all human languages.

This claim of universality has been interpreted in two ways:
(1) every possible phonemic opposition should be provided for (Jakobson's approach), and
(2) every possible phonetic difference should be provided for.

Chomsky and Halle (1968, p. 297) acknowledge the latter level. Ladefoged (1967, p. 57) is of the same opinion: "a linguistic theory should be able to characterize both the oppositions within a language (the difference between the members of the set of all possible sentences) and the contrasts between languages (all and only the features which mark the sounds of the language as being different from the sounds of other languages)."

1.2 Phonological theory should explain why certain sound patterns are favored rather than others, e.g. why five-vowel systems normally have this shape /i e a o u/ rather than, say, /i y ə e ø o/.

* Dept. of Speech Communication, KTH and Dept. of Phonetics, Institute of Linguistics, University of Stockholm, Box 6701, 113 85 Stockholm, Sweden.
1.3 Phonological theory should offer an explanation of the fact that children learn to master the sounds of their language in a specific order (Jakobson, 1969).

1.4 Phonological theory, being a statement about the nature of human language, should include some criteria for defining natural classes of segments. Halle (1964a, p. 328) pointed out that the distinctive feature system is useful for defining these classes, the specification of a natural class involving fewer features than an individual member of such a class.

Furthermore an evaluation procedure is part of the generative theory. An early proposal was described by Halle (1964b, p. 335). The guiding principle was simplicity, in terms of the number of alphabetic symbols. "Given two alternative descriptions of a particular body of data, the description containing fewer such symbols will be regarded as simpler and will, therefore, be preferred over the other."

Halle's suggestion is apparently to base predictions of natural rules, systems, and classes on a measure of phonological distance defined in terms of feature specifications.

2. Articulatory reinterpretation: a discontinuous phenomenon of language acquisition

2.1 The evaluation procedure is not only considered a tool for the linguist, when choosing between competing grammars, but it is also claimed to operate when a child is learning his native tongue (Halle, 1964b, p. 344): "It has been suggested by Chomsky that language acquisition by a child may best be pictured as a process of constructing the simplest (optimal) grammar capable of generating the set of utterances, of which the utterances heard by the child are a representative sample" and similarly (Chomsky and Halle, 1968, p. 251): "An essential feature of our theory of language is that it includes an evaluation measure which makes it possible to assign values to alternative grammars. It is on the basis of this evaluation measure that a child learning a language chooses one of the grammars (of which there are, in principle, infinitely many) compatible with the fairly restricted body of linguistic data to which he has been exposed. The grammar that a child constructs in learning his native tongue will therefore always be the one that ranks highest in terms of this evaluation measure."
In Halle’s view (1964b, p. 344 ff) language change is attributed to discontinuities in the grammars of successive generations (for a criticism of this view, see Weinreich, Labov, Herzog, 1968, p. 143 ff).

2.2 Each child, when learning his native tongue, has to construct his own grammar on the basis of scattered and degraded linguistic data. The primary data available are auditory stimuli, which are given articulatory interpretations by the child. During this process there is a possibility for the child to make an articulatory reinterpretation of perceptually similar segments. Suppose that a given language uses a segment whose phonetic realization is \([\ddot{v}]\). The child, learning this language, is exposed to a number of utterances containing this segment in various contexts and tries to master the pronunciation of it himself. In this particular case, there are more than one articulatory configuration that will produce an acoustic and perceptual pattern similar to that associated with the velarized lateral of the mature speakers. It is for instance possible to use a labio-velar \([w]\) instead. Fig. I-D-1 shows spectrograms of the nonsense words \([\text{iwi}]\) and \([\text{i+i}]\) as spoken by a phonetically trained Swedish speaker. It can be seen that the consonantal portions as well as the transitional patterns are closely similar. It does not seem unjustified to assume that in cases of this sort, that is, in cases of perceptual similarity and articulatory ambiguity, it is more likely than otherwise that the child reinterprets the articulatory characteristics of speech sounds.

This view is not at variance with the view that sound change is grammar change. The point is that the grouping together of certain segments which is not naturally captured in purely articulatory terms can be accounted for if the role of perception is considered. It is not implied that this sound change is gradual. The "sound change" is brought about by a change in grammar, but this change is "suggested" to the child by the perceptually similar effect of two different articulatory gestures.

If it can be shown that segments can be grouped on the basis of perceptual similarity, irrespective of their articulatory characteristics, there may be a strong argument for taking such similarities into account, when constructing a distinctive feature system.

Below we shall demonstrate a few cases where reinterpretations of this sort have occurred and we shall argue that current formulations of phonology
Fig. I-D-1. Spectrograms showing [ili] and [iwi], spoken by a phonetically trained Swedish speaker.
(including marking conventions) do not offer the formal machinery appropriate for predicting reinterpretations in a natural manner.

3. Some cases of articulatory reinterpretation

3.1 The articulatory correlates of the feature flat, introduced by Jakobson et al. (1963, p. 31), are constrictions of the front or back orifices of the vocal tract. Uzbek and Bantu languages, which lack pharyngealized sounds, labialize the sounds that correspond to the Arabic emphatic (pharyngealized) sounds in words borrowed from Arabic, which indicates that there is a great perceptual similarity between pharyngealization and labialization. Jakobson asserts that no language can use an opposition of labialization and pharyngealization, apparently because of the perceptual similarity of the two articulatory gestures. There is nothing in the framework of Chomsky and Halle to suggest such a restriction upon the occurrence of labialization and pharyngealization.

It is noted by Chomsky and Halle (1968, p. 310) that labialization and velarization never act independently in a given phonological system. We suggest that this might be due to the perceptual similarity between the two gestures, a fact that cannot be properly expressed within a purely articulatory system.

On the other hand, the fact that palatalization, velarization, and pharyngealization are mutually exclusive is expressed in a natural way with reference to articulatory factors (Chomsky and Halle, 1968, p. 307).

3.2.1 Latin velars have become labials before dentals in Rumanian, for example /limnum/ → /limn/ and /okto/ → /opt/ (King, 1969, p. 115). This process is expressed as follows in Jakobson’s system:

\[
[+\text{comp}] \rightarrow [-\text{comp}]/\left[\begin{array}{c}
[-\text{cont}] \\
+\text{grave} \\
-\text{comp}
\end{array}\right]
\]

Apparently the process is due to an assimilation, the non-compactness of the dental influencing the velar, while the graveness is retained. If the features are interpreted perceptually, the process is not an unnatural one. In terms of Chomsky and Halle the process is:

\[
\begin{align*}
[+\text{back}] & \rightarrow [-\text{back}] \\
+\text{high} & \rightarrow -\text{high} \\
-\text{ant} & \rightarrow +\text{ant}
\end{align*}
\]
If the process is due to an assimilation, this fact is obscured by the complexity of this formulation of the rule. Another advantage of the Jakobsonian approach in this case is the fact that labials and velars are positively specified by the feature [+grave]. In the Chomsky and Halle approach their common feature is their non-coronality, a feature which they share with many other segments. For a criticism of the specification of labials in the Chomsky and Halle system, see Fant (1969).

3.2.2 The relationship between labials and velars is by no means restricted to Rumanian:

1. Indo-European /kʷ/ corresponds to /kʷ/ in Latin, but to /p/ in Oscean, an Italic dialect (Starkevant, 1961, p. 79), e.g. the equivalent of Latin "quod" is "pod" in Oscean.


3. In the 15th or 16th century there was a change in English [x] → [f], e.g. laugh and cough were pronounced with an [f] instead of an [x] for the spelling -gh (Luick, 1964, p. 623). Similarly, with the spelling -h for [x]: old English dweorh → dwarf (cf. Germ. Zwerg) and beorh → barf (cf. Germ. Berg) (Luick, 1964, p. 1047).

4. In Russian /f/ in loanwords has been actualized as /x/ or /xf/ (Jakobson, 1962, p. 209).

3.3 The feature of "stridency" is in itself a good illustration of the fact that perception is part of the conditions on phonological structure. Stridency is defined by Chomsky and Halle (1968, p. 329): "Strident sounds are marked acoustically by greater noisiness than their nonstrident counterparts. When the air stream passes over a surface, a certain amount of turbulence will be generated depending upon the nature of the surface, the rate of flow, and the angle of incidence. A rougher surface, a faster rate of flow, and an angle of incidence closer to ninety degrees will all contribute to greater stridency." (Italics ours).

It is symptomatic that the definition of stridency is the single case where Chomsky and Halle rely upon acoustical rather than articulatory data. As can be seen from the quotation above stridency can be effected by different articulatory means if only the noisiness, which is a perceptual quality, is preserved.

Chomsky and Halle (1968, p. 177) specify these segments in the following way (only features relevant to the discussion are considered):

<table>
<thead>
<tr>
<th></th>
<th>[f]</th>
<th>[θ]</th>
<th>[s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>coronal</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>strident</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Applying a feature counting procedure, one would rather expect \([θ] \rightarrow [s]\) than \([θ] \rightarrow [f]\), since the former involves a change in one feature and the latter a change in two.

One could improve the situation by making \([f] \rightarrow \text{-strident}\), as suggested by Wooley (1968, p. 141). Then both processes will involve a change in one feature. In this case feature counting will indicate that the processes \([θ] \rightarrow [f]\) and \([θ] \rightarrow [s]\) are equally likely, which is contrary to observed facts. The confusion matrices presented by Miller and Nicely (1955, p. 338 ff) indicate that the distinction between \([f]\) and \([θ]\) is more difficult to perceive than the distinction between \([f]\) or \([θ]\) and \([s]\). Obviously this is due to the fact that both \([θ]\) and \([f]\) are characterized by a low-intensity noise spectrum (Heinz and Stevens, 1961; Stevens, 1968) while \([s]\) is characterized by a high-intensity spectrum.

It appears from the preceding discussion that features have different perceptual weight - in this case stridency prevails over coronality. It seems reasonable to assume that this fact has to be incorporated in the theory, in view of the fact that phonological structure is in part determined by perceptual factors.

3.5.1 The change from an apical \([r]\) to an uvular \([R]\) in German is explained by Zirmunskij (1962, p. 377) in a manner compatible with our definition of articulatory reinterpretation: "Es besteht im Gegenteil aller Grund zu der Annahme, dass der Ausgangspunkt für diesen Vorgang im Deutschen bei der weit verbreiteten Reduktion des alveolaren \(-r\) nach Vokal und im Wortauslaut liegt. Der reduzierte Vokal, der das \(-r\) in dieser Stellung ersetzt, hat artikulatorisch und akustisch eine gewisse Ähnlichkeit mit dem Hinterzungen \(-R\); es führt zum Schnarrlaut bei dem Versuch, den Laut zu verdeutlichen, das heisst ihn deutlicher als Konsonant auszusprechen."
3.5.2 There are two main varieties of /r/ in Swedish. [R] is typical of Southern Swedish, while [r] is typical of Northern Swedish. The uvular variety is clearly gaining ground, the border between the areas slowly moving north-wards. Since there is no trend of migration of Southerners to the North, we must assume that a grammar change of the following kind is spreading north-wards (only features relevant to the discussion are considered):

\[
\begin{align*}
&[+\text{cor}] \quad [-\text{cor}] \\
&[-\text{back}] \quad [+\text{back}] \\
&[+\text{ant}] \rightarrow [-\text{ant}] \\
&\text{/} \\
&[+\text{voc}] \quad [+\text{cons}] \\
&[-\text{low}] \quad [+\text{cont}] \\
&[-\text{high}] \quad [-\text{lat}] \\
\end{align*}
\]

(*) anteriority is involved in the change if we assume an alveolar, rather than a retroflex front /r/.

Despite the complexity of this rule there are many languages, such as French, German, Norwegian, and Swedish, where there is a dialectal variation between a front and a back /r/. It does not seem unjustified to assume that this fact is explicable with reference to perceptual factors. The perceptual similarity of a front and a back /r/ would also offer an explanation to the fact that a phonemic distinction between [r] and [R] is extremely rare. In fact, we know of only one language, Provençal, where [r] and [R] are said to contrast (Coustenoble, 1945, p. 93).

3.5.3 The change /r/ \rightarrow /w/ occurs in Swedish dialects and in English child language (Weir, 1962, p. 49; Leopold, 1947, p. 71; Lewis, 1936, p. 268; Moskowitz, 1970, p. 430). This change is expressed as follows in the Chomsky-Halle framework:

\[
\begin{align*}
&[+\text{voc}] \quad [-\text{voc}] \\
&[+\text{cons}] \quad [-\text{cons}] \\
&[+\text{ant}] \quad [-\text{ant}] \\
&[+\text{cor}] \rightarrow [-\text{cor}] \\
&[-\text{round}] \quad [+\text{round}] \\
&[-\text{high}] \quad [+\text{high}] \\
&[-\text{back}] \quad [+\text{back}] \\
&\text{/} \\
&[+\text{son}] \quad [-\text{lat}] \\
\end{align*}
\]

In view of the great complexity of the rule one would regard this change as quite improbable, but in fact the confusion between /r/ and /w/ is not uncommon at all. We assume that perceptual factors play an important role in this case.
3.6.1 In Polish there has been a change from a velarized lateral [+\text{+}] to a labio-velar semi-vowel [w]. In terms of the Chomsky-Halle framework we get the following rule:

\[
\begin{array}{c|c|c}
+\text{voc} & -\text{voc} & +\text{cont} \\
+\text{cons} & -\text{cons} & +\text{high} \\
+\text{cor} & -\text{cor} & +\text{back} \\
+\text{ant} & -\text{ant} & \\
+\text{lat} & -\text{lat} & +\text{round} \\
-\text{round} & +\text{round} & /
\end{array}
\]

Assuming an evaluation procedure based on feature counting we would have to regard this change as highly improbable. We know, however, that this change is very common, which is quite natural if perceptual similarity is taken into account. Fig. I-D-1 shows spectrograms of the utterances [i+i] and [iwi] as spoken by a phonetically trained Swedish speaker. There are striking similarities between the utterances, as already mentioned in section 2.2.

3.6.2 The same change is reported for English child language (Leopold, 1947, p. 65; Moskowitz, 1970, p. 428).

3.6.3 Latin /l/ was velarized before consonants, according to grammarians of the fourth and fifth century (Bouriez, 1955, p. 255). This velarized lateral has developed into a [w] or [u] in old French, e.g. alba → aube, pulmone → poumon, where the pronunciation was diphthongal, rather than monophthongal as in contemporary French.

3.6.4 The same process has occurred in English as well, in words such as walk, stalk, and talk, where the letters -al- have been pronounced [au] (Luick, 1914, p. 605).

3.7 In the Chomsky-Halle framework the features defining the places of articulation are the same for plosives and nasals. This implies that changes in place of articulation be equally likely among plosives and nasals. Greenberg (1966, p. 65), however, remarks: "that a change m → n is not uncommon, but b → d or p → t is practically unheard of."

Data published by Mohr and Wang (1968, p. 37) indicate that the distances in a perceptual space between nasals are smaller than between the corresponding plosives. The same finding is reported by Pruzansky (1971).
Under the assumption that the probability of a phonologic change is inversely correlated with the perceptual distance between segments, one would have to account for the differences between nasals and plosives in this respect, by including perceptual factors in distinctive feature theory.

The fact that differences in places of articulation are more easily perceived in plosives than in nasals, can probably be explained with reference to the undifferentiated spectral shape of the nasals, which decreases the perceptual value of the place of articulation cues.

Another characteristic of nasals is the fact that they are very often assimilated, as to place of articulation, to a following stop. This fact can be expressed as follows:

\[
\begin{align*}
[\text{+nasal}] \quad & \rightarrow \quad \left[ \begin{array}{c}
\text{\textit{quant}} \\
\text{\textit{bcor}} \\
\text{\textit{\textgamma high}} \\
\text{\textit{\textdelta back}} \\
\text{\textepsilon low}
\end{array} \right] \quad / \quad \left[ \begin{array}{c}
\text{\textit{quant}} \\
\text{\textit{bcor}} \\
\text{\textgamma high} \\
\text{\textdelta back} \\
\text{\textepsilon low} \\
\text{\textepsilon\text{-cont}}
\end{array} \right]
\end{align*}
\]

This is descriptively correct, but in the proposal of Chomsky and Halle there is no particular reason why this process should not as easily affect [\textit{-nasal}] segments. Again we suggest that the explanation of the behavior of nasals, as opposed to other segments, is to be found in the small perceptual value of place of articulation distinctions. In the position before a stop the articulatory effort of maintaining a separate place of articulation of the nasal is rewarded by a very small perceptual effect. So the "lazy mouth" is allowed to minimize its articulatory effort by assimilating the place of articulation of the nasal to the following stop.

4.

To what extent will current phonology be capable of handling articulatory reinterpretation? It might for example be proposed that marking conventions be extended so as to include reinterpretation. In our opinion it might be possible to do so but such modifications are not likely to eliminate the descriptive nature of the conventions. For a discussion of the intrinsic content of features, see Fromkin (1968).

To take one example, by convention no. XI is stated that if a language has five vowels, the natural case is a vowel system of this kind i e a o u; (2) e a o u.
a system such as \( \hat{a} \) is filtered away, although system (4) was not dis-

\( \hat{\Lambda} \)

\( \omega \) tinguished from (2) by the evaluation measure. What is the basis for regard-
ing (2) as more natural than (4)? Chomsky and Halle, in fact, do not answer this question (1968, p. 401): "Thus, although a vowel system such as (2) would be more natural, in some significant sense, than one such as (3) or (4), our measures of evaluation make no distinction among them." Since no other explanation is offered we must assume that the intuition of Chomsky and Halle is based upon frequency of occurrence. This approach is of course not a new one, e.g. in the Principles of the International Phonetic Association (1963, p. 5 ff) the front unrounded and back rounded vowels are called primary, while the front rounded and back unrounded ones are called secondary.

It would seem then that the marking conventions primarily serve the pur-
pose of stimulating interest in a more precise formulation of the goals of phonological theory. As such they represent an invaluable contribution to linguistic thinking. The concrete proposal of Chomsky and Halle is however subject to criticism once these goals and the requirements on explanations have been set up.

Wang (1968, p. 45) tried to explain the shape of simple vowel systems with reference to a principle of maximizing the perceptual contrasts between segments. Fromkin (1968, p. 164) remarks: "We shall probably find that here [to explain the shape of vowel systems J.J.] perceptual properties are more important than articulatory ones."

The explanatory value of the contrast principle has been further explored by Liljencrants and Lindblom (forthcoming). Results based on numerical computation indicate that this principle definitely plays a role, although other factors are involved too.

It appears from the preceding discussion that it is necessary to consider the intrinsic content of segments. Jakobson's approach to distinctive fea-
tures in many respects satisfies the perceptual requirements of "intrinsical-
ness".

(Jakobson et al., 1963, pp. 12 and 13): "The closer we are in our investi-
gation to the destination of the message (i.e. its perception by the receiver), the more accurately can we gage the information conveyed by its sound shape.
This determines the operational hierarchy of levels of decreasing pertinence: perceptual, aural, acoustical, and articulatory (the latter carrying no direct information to the receiver)."

"To sum up, the specification of the phonemic oppositions may be made with respect to any stage of the speech event from articulation to perception and decoding, on the sole condition that the variables of any antecedent stage be selected and correlated in terms of the subsequent stages, given the evident fact that we speak to be heard in order to be understood."

In fact, speakers of Bantu languages, when pronouncing Arabic loan-words, behave in accordance with the perception-oriented definition of the flatness feature, as already mentioned in section 3.1.

Since the child's first contact with his mother-tongue is mediated by auditory, not articulatory stimuli, it should not cause surprise that perceptual factors shape phonological structure. Consequently there is no reason to expect that a theory based principally on articulation (Chomsky-Halle) should give an adequate description of the events.

On the other hand, the price paid for Jakobson's listener-oriented approach is that complications arise in the assignment of a correct articulatory interpretation of flatness, as pointed out by McCawley (1967, p. 112 ff).

Rather, it seems reasonable to expect that all links in the speech communication chain should contribute to phonological structure, e.g. articulation favors and disfavors certain segments, while perception involves other preferences and constraints. Consequently we can only hope to explain phonological rules and phonological structure if we succeed in constructing adequate models of all aspects of the speech communication process. For a discussion of some concrete applications of phonetic theory to phonology, see Jonasson (forthcoming), Lindblom (forthcoming), and McAllister (forthcoming).

References:
(The date of publication refers to the editions used by the author. The original date of publication, when known, is indicated in square brackets)


McCawley, J. D. (1967): "Le rôle d'un système de traits phonologiques dans une théorie du langage", Languages 32, pp. 112-123.


