ordic Journal of Linguistics

ntinuation of the *Norwegian Journal of Linguistics/Norsk tidskrift for sprogvidenskap.*

NSORS:
Nordic Humanities Research Councils

TORS:
or: Fred Karlsson, Helsinki
ditor: Marja Leinonen, Helsinki

TORIAL BOARD:
Allén, Göteborg
Basbøll, Odense
nn Benediktsson, Reykjavik
t Fossestøl, Oslo
stein Fretheim, Trondheim
Hakulinen, Helsinki
o Korhonen, Helsinki
Elisabeth Ingram, Trondheim
Björn Lindblom, Stockholm
Per Linell, Uppsala
Ragnar Rommetveit, Oslo
Ulf Teleman, Roskilde
Ole Togeby, København

TORIAL POLICY:
Nordic Journal of Linguistics is the journal of the *Nordic Association of Linguists.* ournal is primarily devoted to the publication of contributions from the members of ssociation, but contributions from non-members may be submitted to the Editorial for consideration. The Nordic Journal of Linguistics is concerned with all hes of linguistics. Preference is given to contributions of general theoretical or odological interest and to studies on Scandinavian languages (including Finnish and ish). In addition to original articles the journal will publish replies and comments on es published in the journal and review articles or brief reviews of current books. The al will be published in two issues annually, forming together one volume of ximately 200 pages.

NUSCRIPT ADDRESS:
nuscripts, conforming to the rules on page 3 of this cover, should be addressed to ssor Fred Karlsson, Department of General Linguistics, University of Helsinki, uskatu 11-13, SF-00100 Helsinki 10, Finland.

offprints of each article and review article and ten offprints of shorter reviews will pplied free of charge, and additional offprints can be ordered.

NESS ADDRESS:
nunications regarding accepted manuscripts, proofs, orders for offprints, subscrip- etc. should be sent to the publisher, Universitetsforlaget.

LISHER:
rsitetsforlaget, Journals Dept., P. O. Box 2959, Tøyen, Oslo 6, Norway.
can office: P. O. Box 58, Irvington-on-Hudson, New York 10533, USA.

CRIPTION PRICE: NOK 160.00 USD 28.00.
E PER ISSUE: NOK 100.00 USD 17.00.

d by Fr. Bagges kgl. Hofbogtrykkeri, København.

Palatal Vowel Harmony: A Perceptually Motivated Phenomenon?

Kari Suomi

Suomi, K. 1983: Palatal Vowel Harmony: A Perceptually Motivated Phenomenon? *Nordic Journal of Linguistics* 6, 1-35.

The paper attempts to determine the motivating causes of palatal vowel harmony (PVH). Previously suggested causal explanations of PVH are critically evaluated, especially the progressive palatal assimilation view of the origin of PVH. Data on PVH restrictions from Finnish and Turkish are examined against a set of perceptually motivated working tendencies. It is shown that PVH and "labial harmony" are only special cases of a single, unitary type of restriction, statable acoustically in terms of the frequency of the second formant. A perceptual theory of the causes of PVH is proposed and general phonetic conditions for and against the development of PVH in a language are suggested. On the whole, the paper is an argument for a substance-based approach to phonology.

Kari Suomi, University of Turku, Department of Finnish and General Linguistics/Phonetics, Henrikinkatu 2, 20500 TURKU 50, Finland.

0. INTRODUCTION

Palatal or vertical vowel harmony, as distinct from other types of vowel harmony which will not be discussed in this paper (see e.g. Anderson 1980 for an account of the formal properties of all types of vowel harmony), is the phenomenon common in Altaic and Uralic languages whereby the selection of vowels in the non-initial syllables of words is constrained by the vowel in the first syllable. A common formulation of the main restriction of palatal vowel harmony (henceforth PVH) is that all vowels within a word must be either *front* or *back*, i.e. they must agree with respect to *palatality*. Previously suggested explanations of the causes of PVH can be subsumed under two broad classes. In the first class of proposals a supposedly universal, physiologically based principle of saving muscular effort is invoked to account for the non-combinability of *front* and *back* vowels in the same word in PVH languages, this involving (it is claimed) a motivating saving in terms of the movements of the organs of speech (see e.g. Hakulinen 1968:18; Lewis 1967:15-16). Let me point out just two obvious objections against this type of explanation, one based on general typological considerations and the

other on a circumstance obtaining in PVH languages themselves. Firstly, then, there is the massive counter-evidence from the non-PVH languages of the world in which the combination of *front* and *back* vowels in the same word causes no problems. The claim that PVH is caused by a physiologically determined universal tendency to avoid extreme articulatory movements within the word immediately raises the question why the tendency obviously fails to operate for the great majority of languages. Clearly, a reasonable null hypothesis would state that a universal tendency has similar effects across the board. It is possible, of course, that a true universal tendency is only triggered (or suppressed) under some further specific conditions, but then these conditions should be carefully determined. In the case of the present alleged universal tendency, no such motivating or inhibitory conditions have been offered to explain the difference between PVH and non-PVH languages. It seems to me that even the weak version of the economy of effort hypothesis maintaining a *propensity* towards avoiding extreme articulatory movements (as against a *necessity* claimed by the strong version, see Lindblom 1980) cannot provide a consistent solution to the dilemma. This is not to deny the reality of the economy of effort principle as such, only to argue that it cannot be syntagmatically extended to account for the co-occurrence restrictions in PVH languages in the light of the counter-evidence (while it is still possible that PVH also brings about an amount of saving in terms of articulatory movements). Secondly, speakers of PVH languages find it perfectly easy to pronounce non-harmonious sequences of vowels across word boundaries (and in polymorphemically interpreted compound words) even at very great rates while they have difficulties with pronouncing similar sequences of vowels separated by one or more consonants as long as the vowels belong to the same word (e.g. in loanwords critically violating the native PVH restrictions). The word is notoriously not a well-definable articulatory unit in normal connected speech, nor do word boundaries usually involve abrupt breaks in the flow of articulatory movements, and hence the ease with which non-harmonious vowel sequences containing a word boundary are pronounced cannot be attributed to any clear-cut articulatory difference between such sequences and those contained within a single word. The conclusion is that the difficulties encountered with non-harmonious sequences not containing a word boundary (or a transparent morpheme boundary) are not due to constraints on articulatory movements (in the sense of genuine language-independent limits on possible articulatory gestures), and while the amount of muscular energy that has to be expended in speech production (alias the principle of least effort, or that of "minimal articulatory antagonism", Lindblom and Sundberg 1971) may well operate in shaping human language, it cannot be accepted as the cause of PVH. Additional difficulties connected with the articulatory explanation, including a particular variety known as the theory of palatal attraction, will be pointed out below.

The second class of suggested explanations is based on an explicit recognition of the central position of the word in PVH systems. Here PVH is regarded as being caused by the function it serves to fulfill. The function can be called a *delimitative* one, PVH acting as a word boundary signal (an "aphonematisches Grenzsignal" in the terminology of Trubetzkoy 1971:250-251). This explanation cannot be completely rejected since it is true that a change from one harmonic class to another within an utterance signals the presence of a word boundary somewhere between the non-harmonious vowels, and that PVH thus may contribute along with the other cues to delimiting successive words from each other. However, PVH can at best function in this capacity to a limited extent only, since it can provide only a *positive* cue in the case of a shift from one harmonic class to another at some point of an utterance. Thus, PVH alone can never inform the hearer that a word boundary has *not* occurred, nor can it contribute to locating existing word boundaries between successive words belonging to the same harmonic class. An argument that reduces the attractiveness of the Grenzsignal explanation, then, derives from the limited power potentially attached to the delimitative function of PVH, from the conspicuous disproportion between the function and its implementation. On the one hand we have the word boundary signalling function performed only occasionally (potentially only between hetero-harmonious words) by a device not made use of in the majority of languages, and on the other hand there are what seem to be too high a price for this modest achievement, e.g. the severe restrictions on the number of possible words for the purposes of word formation (as the paradigmatic resources of vowel distinctions are constrained by the syntagmatic rules for combinability), the resulting complexity of derivational and inflectional morphology by way of semantically empty alternations, etc. Nevertheless, feeling unable to reject fully the Grenzsignal explanation of the cause of PVH and acknowledging that single cause theories seldom tell the whole truth, I must leave it to the reader to weigh its relative merits against those of the perceptual theory presented below.

A serious shortcoming common to both classes of explanations discussed above is that they leave us very much in the dark as to the question of what sort of circumstances might favour the development of PVH. In other words, while most (all?) languages have *front* and *back* vowels, why is PVH so rare? I will attempt to answer this question in the final section.

PVH has been extensively discussed in phonological theories. In particular, a wealth of generative descriptions of PVH systems is available. However, I have not come across explicit proposals of the causes of PVH in this literature except for Wiik's remark that, if generative phonogy is conceived of as being "mentalistic", then it pays main attention to the energy expended by the brain (Wiik 1975:42), implying that vowels defined as *marked* in the framework demand more processing by the brain than *unmarked* ones. If this is taken to be a suggestion for an explanation of the causes of PVH then it of course competes with the others as a possible candidate, and it has to be similarly evaluated against pertinent data (as a simple proclamation cannot make a descriptive framework psychologically realistic). More specifically, it needs to be shown that the formal operations postulated by generative phonology do indeed have their counterparts in the operations executed by the brain. Until then this particular descriptive framework remains just one (albeit highly rigorous) formalization of the linguistic product out of an infinite number of alternative possibilities (and this is true of all descriptive frameworks detached from the actual processes taking place in the use of language). Generative phonologists themselves do not usually even attempt to produce experimental evidence in support of their contentions, typically regarding such information as irrelevant (or, at most, as pertaining to 'mere' performance). As it stands, the case for the psychological reality of standard generative phonology seems very weak (for a thorough assessment, see Linell 1979).

Contemporary descriptions of PVH typically operate with (notational variants of) traditional articulatorily defined vowel features such as *low-high, front-back* and *unrounded-rounded* (which, according to Anderson 1980:7, "have independent phonetic motivation and validity"). However, investigations of the positions and movements of the relevant articulatory structures during speech production have made it clear that the terms cannot be given a straightforward literal interpretation because of many discrepancies between the classificatory use and articulatory fact. In contrast, it has been established that there is a more systematic relationship between the

acoustic characteristics of vowels and their classificatory properties (while the traditional terms may still be retained as pseudo-articulatory labels in fact referring to the acoustic characteristics: this practice is also followed in this paper). The standard specification of the acoustic characteristics of vowels is in terms of the centre frequencies of the two or three lowest energy maxima or formants (see e.g. Lindblom and Sundberg 1971). Recently Nearey (1980) has argued for a two-dimensional specification of basic vowel qualities (excluding e.g. the "r-colored" vowels of American English) on perceptual grounds, noting that the adoption of a two dimensional analysis has important implications for phonology. It is likely that the perceptually relevant dimensions are F1 and F2', the latter being a weighed average of F2 and F3 (and possibly of higher formants). The F2' concept has been introduced by Fant (see e.g. Fant 1973) to obtain a closer correspondence between acoustic measurements and the perception of vowel quality, especially in the region of *high front* vowels. However, while the perceptual reality of the F2' concept seems unquestionable (Carlson et al. 1970) there are problems connected with its precise calculation from e.g. spectrographic data, and no great harm should be done in the present context if the acoustically more straightforward dimension of F2 is used instead. To return to Nearey's stimulating comments, he (Nearey 1980:239) specifically questions the wisdom of treating *lip rounding* as an independent dimension of vowel quality because, like the traditional *front-back* dimension, this feature mainly affects the position of F2 (or F2'), *roundedness* in a clearly definable way reinforcing the perceptual *backness* of *back* vowels and *unroundedness* correspondingly reinforcing the perceptual *frontness* of *front* vowels. Even more pertinently, Nearey further points out the dependence of "rounding harmony" in PVH language on "advancement harmony", and the difficulties connected with stating the harmonic classes of Finnish in terms of the traditional three-dimensional articulatory system, and the ease with which this can be done in terms of F2 (Nearey 1980:239).

A conventional two-dimensional specification of vowel quality in terms of F1 and F2 is also adopted in the present paper as a first approximation to the crude perceptual dimensions involved. It is likely that the uncertainties concerning the link between the details of the acoustic signal and their perception by the hearer that still persist become tangible only in areas which from the point of view of PVH can be considered peripheral, and that a whole-spectrum approach

instead of the formant-based one adopted would entail very similar consequences in this respect (cf. Bladon and Lindblom 1981). However, the precise definition and quantification of the concept of perceptual distance, essential to the theory proposed in this paper, must remain to be performed in the future.

The distinction between typologically primary and secondary vowels is of prime importance in the description of PVH. The former, as is well known, include all acoustically extreme vowels (in an F1-F2 plot) except the *low front* [æ] which, together with all acoustically interior vowels (see Crothers 1978), belongs to the secondary vowels. It is generally agreed that the primary vowels are typologically more common because they form the set of vowels that are perceptually maximally distinct from each other. To reflect this causal relationship correctly (which I will take for granted in the following), I will use the terms *strong* and *weak* for the primary and secondary vowels, respectively, thus referring directly to the ultimate phonetic causes of the empirically established typological differences.

It is interesting to observe that the vowels generated by the Lindblom/Sundberg articulatory model (Lindblom and Sundberg 1971) under the condition that all parameters except the degree of jaw opening (i.e. lip shape, the position of the tongue tip, the shape and position of the tongue body, and larynx height) are set to zero come very close to the *front* weak vowels, ranging as they do approximately from [æ] to [ü] (see Lindblom 1972, esp. Fig. 5 on p. 74). To the extent that vowels with such an articulatorily neutral manner of production are characterized by a maximum amount of articulatory ease (or, conversely, minimum amount of articulatory effort), my interpretation of the situation is that considerations of articulatory effort involved in such simple vowels have at most negligible effects on what sort of vowels are most favoured in the sound patterns of the languages of the world. In other words, it seems to me that in the case of simple oral vowels the demands for sufficient perceptual contrast by far overrule those of articulatory ease.

The main thesis of the present paper is that PVH is a perceptually motivated phenomenon, prompted by the occurrence of weak vowels in the non-initial syllables of the word. The latter circumstance would, as such, necessitate rather precise F2 computing in the non-initial syllables in addition to the initial one, weak vowels typically having nearby strong vowel neighbours on the same F1 level. In PVH languages, however, the need for accurate vowel quality judgments based primarily on differences in F2 is effectively reduced after the introductory, initial syllable of the word. This is accomplished by developing contextual restrictions, sensitive to the vowel of the initial syllable, on the occurrence of vowels in the subsequent syllables. These restrictions, the PVH rules of combinability, limit the choice of possible vowels differing along F2 in the non-initial syllables after a given vowel in the first syllable and thus, by excluding some paradigmatically possible vowels as syntagmatically impossible or ungrammatical, make the perceptual decision in the non-initial syllables easier for the hearer.

A more detailed and substantiated account of the suggested perceptual theory of the causes of PVH will be given in section 3 after presentation of some relevant data from (the history of) Finnish and from Turkish in the immediately following sections. However, rudiments of the theory are necessary for a meaningful discussion of the data, and these are given below in the form of a number of perceptually motivated working tendencies to be evaluated against the data. However, let me first define a further distinction between *unbounded* and *bounded* strong vowels. By unbounded strong vowels I mean strong vowels which do not have (neighbouring) weak vowels on the same F1 level (sharing the same specification for *height*) in the syllable position in question, and by bounded strong vowels those strong vowels that do have at least one (neighbouring) weak vowel with (approximately) the same F1 in the syllable position in question. For example, /e/ is the only *mid* vowel in Turkish capable of occurring in the non-initial syllables of words (irrespective of what vowel may occur in the first syllable) and is thus an unbounded (strong) vowel by our definition; on the other hand, three *mid* vowels are capable of occurring in the initial syllable, viz. /e/, /ö/ and /o/, and both /e/ and /o/ would count as bounded vowels in this syllable position (if indeed the distinction were relevant in the first syllable, see below).

PVH is a "left-to-right" co-occurrence restriction, and the working tendencies are most suitably formulated as constraints on the occurrence of vowels in the non-initial syllables. In the following sections the Tendencies will be frequently referred to with the numbering given below:

Tendency (1): The occurrence of weak vowels in non-initial syllables tends to be restricted in such a way that, in principle, F1 judgment is sufficient for their correct identification, given the prior correct identification of the vowel in the first syllable. In

other words, there is a tendency to make the identification of non-initial weak vowels predictable from the vowel in the first syllable and information on the (approximate) F1 frequency of the non-initial weak vowels themselves. The resulting elimination of the need for accurate F2 computing is the motivation for the restrictions imposed on the weak vowels.

Tendency (2): Bounded strong vowels tend to be subject to PVH restrictions in order to ensure the predictability of the neighbouring weak vowel on the basis of the vowel in the first syllable. Thus, since weak vowels are perceptually less salient than strong vowels, their presence in the non-initial syllables must be signalled by particular vowels in the first syllable, and these in turn may not be allowed to be followed by the neighbouring bounded strong vowel. In this way the weak vowel(s) and bounded strong vowel(s) on the same F1 level, differing among themselves along F2, enter into harmonic alternation by grouping into mutually exclusive sets.

Tendency (3): Unbounded strong vowels tend to be free of PVH restrictions, i.e. occur after any vowel in the first syllable. Since unbounded strong vowels do not have weak neighbouring vowels on the same F1 level, the need for accurate F2 computing (unsupported by a F1 difference) cannot arise, and hence there is no perceptual motivation for restricting the occurrence of these vowels (strong vowels on the same F1 level always being widely separated from each other along F2 or, in the case of [a], alone on the F1 level).

1. THE DEVELOPMENT OF PVH FROM PROTO—URALIC TO MODERN FINNISH

The historical perspective adopted in this section is necessary for the illustration of the workings of the perceptual tendencies, especially as they are reflected in the behaviour of the "neutral" vowels. The present situation cannot be understood without some knowledge of the past stages. Needless to say, the reconstructions referred to below are hypothetical constructs inevitably involving an amount of uncertainty.

The discussion of PVH in this section will take place in terms of disyllabic word types because, ever since Proto-Uralic, uninflected words have in the overwhelming majority of cases been disyllabic,

and the harmonic relationships in the subsequent syllables have been completely determined by the vowel in the first syllable (for some fluctuation in this respect in Modern Finnish after initial syllables with only "neutral" vowels, see below). Therefore, only the vocalism of the first and second syllable will be given in the form /vCv/ in which *C* stands for any consonant (combination) and is meant to contain the syllable boundary. For the most part I will be dealing with the /single/ (phonetically short) vowels only. This is a legitimate simplification because the harmonic relations are practically identical among the /single/ and /double/ monophthongs (and the diphthongs which structurally count as /double/ vowels, homosyllabic sequences of vowels behaving with respect to PVH as if there were a syllable boundary in the middle). I will write the vowels between the slashes everywhere, regardless of whether they should be considered full phonemes or not. The asterisk (*) denotes reconstructed forms and vowels. For a "family tree" of the Uralic languages, with approximate estimates of the ages of the various reconstructed periods, see Häkkinen (1981).

The vowel systems to be discussed will be presented in a schematized form of the conventional acoustic F1-F2 arrangement along quasi-logarithmic scales (exact Hz values are not given for obvious reasons but the scale determines the overall shape of the arrangement). For the vowels occurring in Modern Finnish the positions indicated are approximately those valid for the /single/ monophthongs (for formant data see e.g. Wiik 1965), for others the positions are rough estimates based on articulatory information contained in historical accounts and on general considerations of the acoustic consequences of the major articulatory dimensions (see e.g. Lindblom and Sundberg 1971, Crothers 1978, Nearey 1980). A reference arrangement of the relevant vowels is shown in Figure 1 to introduce the transcriptional conventions used subsequently (on the whole typographical ease has determined the choice of the transcripts). The enclosed area in Figure 1 encompasses the weak vowels: /ü, ö, å, o, u/ are *rounded* and the others *unrounded*.

According to the latest reconstruction (Janhunen 1981, see also Sammallahti 1979, 1980) the /short/ vocalisms of the first two syllables of Proto-Uralic were as shown in Figure 2, in which the left hand side refers to the first and the right hand one to the second syllable. The vertical line in both syllable positions shows the division of the vowels into mutually exclusive harmonic sets.

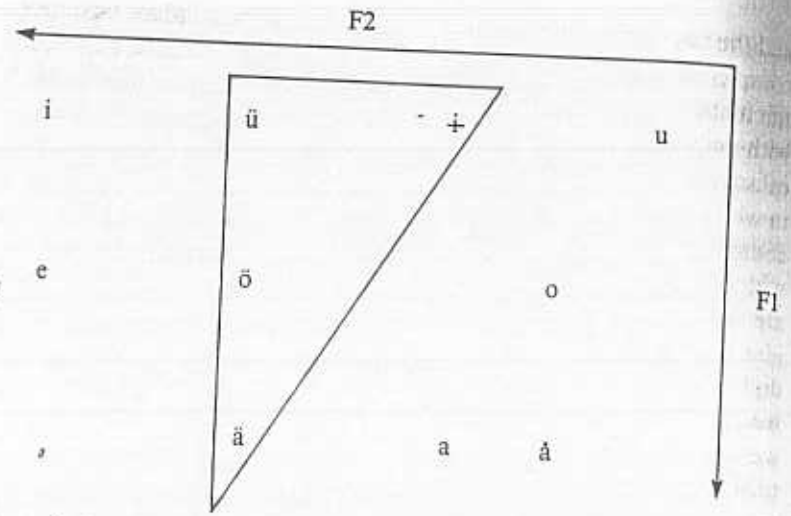

Figure 1. A schematic arrangement of the relevant vowels in a conventional acoustic F1 — F2 diagram. The dimensions have logarithmic scales. For further explanations see the text.

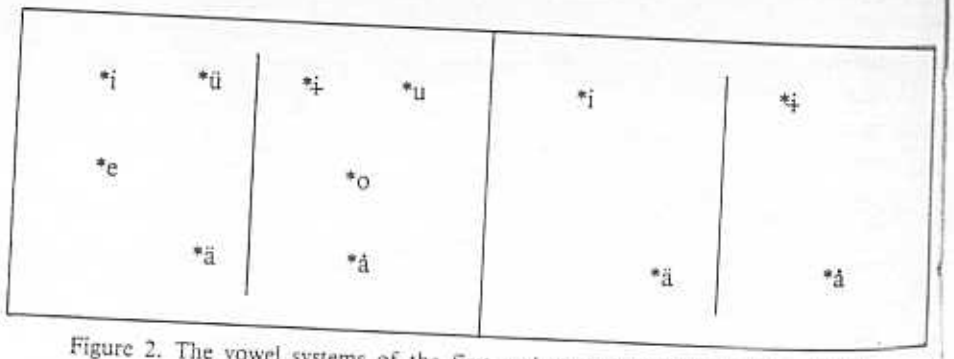

Figure 2. The vowel systems of the first and second syllables in Proto-Uralic words. For further explanations see the text.

During Proto-Uralic, then, only *front* vowels could occur in a word with a *front* vowel in the initial syllable, and only *back* vowels could follow an initial *back* vowel. At this stage it indeed seems as if *palatality* vs. *non-palatality* (as an articulatory dimension) were the part and parcel of PVH. However, let me briefly comment on the situation in the light of the working Tendencies put forth above (using, as was agreed, the corresponding numbering):

(1) As predicted by the first Tendency, the weak vowels in the second syllable are restricted in their occurrence in such a way that they are fully predictable from the vowel in the first syllable and information on their own F1. At this stage in fact also the strong vowels in the second syllable are similarly predictable (only one vowel with a given F1 being permitted after a given initial vowel). As will be seen below, the predictability later became less powerful with the introduction of more vowels into the second syllable. Nevertheless, the situation shown in Figure 2 clearly indicates that PVH effectively eliminates the need for accurate F2 computing in the non-initial syllables.

(2) The bounded strong vowels */i/ and */å/ are subject to PVH restrictions as predicted by the second Tendency, occurring in mutually exclusive environment with their respective neighbouring weak vowels */ɨ/ and */ä/, thus forming pairs of vowels in harmonic alternation. Tendency (2) contains the claim that the restrictions imposed on bounded strong vowels are not due to these vowels themselves but are caused, somewhat indirectly, by the demand that the neighbouring weak vowel maintain its predictability (and this can of course be achieved only if the vowels adjacent along F2 occur in mutually exclusive environments). The claim gets its first bit of support from what is reported to have happened to the non-initial *high* vowels of Proto-Uralic by the time of Proto-Finno-Permic. According to Janhunen's reconstruction */ɨ/ changed to */i/ during this interval (cf. below), the result being that the latter, now the only *high* vowel in the second syllable, became "neutral", i.e. capable of occurring after both *front* and *back* vowels (Janhunen 1981:249). Janhunen (ibid.) interprets the development as a loss of PVH for the *high* vowels. This is superficially true, of course, and the new situation is difficult to handle for anyone thinking that the essence of PVH is the requirement that all vowels within a word agree with respect to *palatality* (I do not insinuate that Janhunen holds such views; I am simply trying to argue against the articulatorily based explanation of PVH). However, Janhunen's interpretation gives no clue as to why PVH suddenly ceased to operate for one vowel while it still influenced the others (from now on */i/ in the second syllable could follow any vowel in the first while the old restrictions remained for the *low* vowels, see below). In the present perceptual framework, in contrast, PVH did not cease to operate, only the general motivations for its application were lost. The loss of */ɨ/ changed */i/ from bounded to unbounded, the latter being now perceptually very safe alone on its F1 level. Being unbounded by nearby weak vowels on the same F1 level, the new behaviour of */i/ is in perfect agreement with the predictions of Tendency (3). PVH is motivated

by the presence of weak vowels in the non-initial syllables and when they are absent (or disappear) there is no need for the restrictions. More confirmation of the independence of strong vowels, as such, from PVH restrictions will be seen below.

(3) There were no unbounded strong vowels during Proto-Uralic (cf. immediately above).

By roughly early Pre-Finnic, several changes relevant to the development of PVH had taken place in the system (in naming the various historical periods I attempt to follow Sammallahti 1980). Thus */å/ and */i̯/ disappeared during this interval, the former changing to */a/ everywhere and the latter to */a/ (and thus merging with earlier */å/) in the first syllable and to */i/ (merging with old */i̯/) in the second (Janhunen 1981:248, Sammallahti 1980:4-5). Later, */i/ in the second syllable changed to */e/, and */o/ was introduced as a new vowel into the second syllable. Two novel word types were introduced, viz. */iCa/ and */eCa/ in addition to the old */iCä/ and */eCä/ types, both */i/ and */e/ thus becoming "neutral" in the first syllable. The system at approximately Early Pre-Finnic is shown in Figure 3, the circles indicating "neutral" vowels.

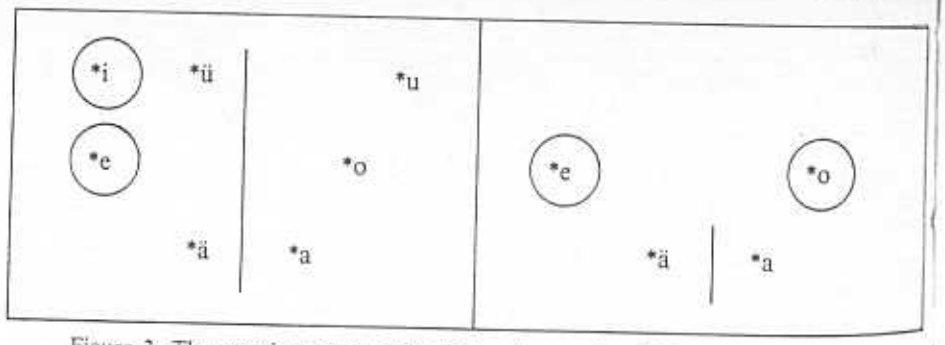

Figure 3. The vowel systems of the first and second syllables in approximately Early Pre-Finnic words. For further explanations see the text.

During this period */i/ and */e/ in the first syllable could be followed by any of the vowels in the second syllable, and */e/ and */o/ in the second syllable could occur after any vowel in the initial syllable; the rest of the vowels in both syllable positions were divided into two harmonic classes as indicated by the vertical lines in Figure 3. I will again comment on the new situation with reference to the numbered Tendencies, discussing first the developments leading to and following the "neutralization" of */i/ and */e/:

(1) After the emergence of the new word types */iCa/ and */eCa in addition to the old */iCä/ and */eCä/, the first Tendency was n longer obeyed in that the non-initial weak vowel */ä/ could now oc cur in the same environment as */a/ and was thus not predictable a the only permitted *low* vowel after initial */i/ and */e/ (as had pre viously been the case). Tendency (2) was at the same time disobeyed quite trivially, in that the bounded strong vowel */a/ was no longe subject to the previous PVH restrictions when following the newly "neutralized" vowels; notice, however, that */a/ continued to be prohibited after the other *front* vowels (i.e. */ü/ and */ä/, later also */ö/). Now the question crucial to the validity of the first Tendency is this: what happened, if anything, to restore the earlier predictability of the weak */ä/ in the second syllable after the now "neutral" vowels in the first syllable?

Two important developments are reported to have taken place, albeit only gradually. On the one hand, the old */iCä/, */eCä/ types started slowly to become unproductive as patterns of novel word formations, to the extent that in Modern Finnish new words with /i/ and /e/ (or their combinations in diphthongs) in the first syllable hardly ever take /ä/ in the second syllable, the productive pattern of "neutral" vowel + *low* vowel being /iCa/ and /eCa/. The only exceptions are completely transparent derivatives of the old types used primarily to coin new proper names (Itkonen 1948). The other important development, then, is the rise in productivity of the new types. The circumstances relevant to the present discussions are as follows.

It is generally agreed that */iCa/ was introduced earlier than */eCa/, i.e. */i/ became "neutral" before */e/ (the former was such already during Proto-Finno-Ugric, see e.g. Sammallahti 1980:4). Further, the investigations by Itkonen (1948) of the productivity of various word types during approximately Finno-Ugric on the one hand and during Finno-Permic and later times on the other, strongly suggest that */iCä/ lost its productivity before */eCä/ and that */iCa/ rose in productivity before */eCa/ (the figures given by Itkonen, to the extent that they are representative of the word types as they have really existed, are quite conclusive in this respect). Moreover, taking into account what is now considered likely concerning the approximate time of the introduction of the new */iCa/ and */eCa/ types into the language, it is highly probable that */iCä/ and */eCä/ started to become unproductive only after the introduction of the new types. So much, but not much more, can be inferred

from the information available. Even so, I would find it hard to believe that the parallelism of the fates of */i/ and */e/ is a mere coincidence. Instead, the data strongly suggest that the coming of */iCa/ and later of */eCa/ were the direct causes of the growing unproductivity of */iCä/ and */eCä/. In the present perceptual framework these developments can be given a very consistent and straightforward explanation: in order to restore the predictability of the weak */ä/ after all initial vowels, a new restriction was slowly emerging to the effect that */ä/ was no longer permitted to occur after */i/ and */e/ since */a/, differing from */ä/ with respect to a critical F2 distance only, also occurred after these vowels. Because of the new restriction, the old types */iCä/ and */eCä/ were no longer so easily available for the purpose of forming new words. In this weak sense, then, they became synchronically ungrammatical. However, the old combinations may have been very common in the vocabulary, and the new restriction was not strong enough to make the existing words obsolete; in fact the old types seem to have remained decreasingly productive long after the arrival of the new, conflicting types (Itkonen 1948:141-142). This is not a very strong argument against Tendency (1) and the perceptual theory in general since it is based, quite explicitly, on perceptual *tendencies,* not on absolute limits on perceptibility. That Tendency (1) is at work, however, is shown by the growing (and now near enough total) unproductivity of the old combinations. Otherwise it would be hard to account for these developments (cf. Itkonen 1948:140-141).

As the old */iCä/, */eCä/ combinations did not disappear, words in this pattern still exist in Modern Finnish as unproductive fossilizations reflecting the permitted combinations of an earlier system. These include non-decomposable nominals as well as originally derived verbs and function words (see e.g. Groundstroem 1971:112-114). Another remnant of the old system is the fact that words containing only "neutral" vowels in the uninflected form (nominative singular) take weak harmonic (i.e. *front*) vowels in inflectional suffixes (e.g. *tie + tä,* partitive sg. of *tie* 'road', *vene + ttä* 'boat', *enkeli + ä* 'angel' etc.). In what follows I will discuss, except when explicit mention to the contrary is made, only the development of the PVH restrictions as they pertain to the productive disyllabic words. Notice that with regard to all other initial vowels than /i/ and /e/ there has never been any fluctuation as to the harmonic class of the subsequent vowels except for quite recent, non-nativized loanwords violating the native restrictions (these cases, exhibiting vacillation even for a single speaker, deserve a separate treatment).

I have not yet discussed the origin of the new word types that brought about such far-reaching changes in the system. The sources available to me suggest that the later */eCa/ was borrowed (Sammallahti 1980:5), possibly from the Baltic or Germanic languages (Itkonen 1948:140), but I have not come across any indication of the probable origin of the older */iCa/ type. Although it would be tempting to guess that this type was also borrowed (so as to make the parallelism perfect) this can hardly be considered legitimate in the absence of any independent indications in that direction. Notice in passing that there may also be some system-internal reasons for the triumph of the new types over the old ones and for the ensuing quasi-"neutrality" of /i/ and /e/. As Finnish makes use of three linguistically relevant degrees of F2, the middle one being occupied by the weak vowels (for a more explicit acoustic description of the Finnish vowels see towards the end of the present section), it may be natural that one of the sets of strong vowels so to speak falls outside the system of harmonic alternation, not belonging to either of the harmonic sets and being only subject to the paramount constraint that weak vowels in non-initial syllables have to be predictable. Whatever the reason, observe that from now on weak vowels in the non-initial syllables of words belonging to the productive vocabulary were signalled by only weak vowels in the first syllable, /i/ and /e/ now signalling, like the *back* strong vowels, only the absence of weak vowels in the subsequent syllables.

(3) As predicted by the third Tendency, the unbounded strong vowels in Figure 3 are free of PVH restrictions. This is another bit of support for the claim that eventual PVH restrictions on strong vowels are caused by the presence of neighbouring weak vowels: at this stage there was no *mid* weak vowel in the second syllable and hence no (need for) restrictions on the *mid* (unbounded) strong vowels. The combinatory possibilities of second syllable */e/ and */o/ during this period also indicate once more that what is called PVH has obviously nothing to do with an alleged physiologically determined need to avoid the combination of *front* and *back* vowels in the same word.

Later on */i/, */ü/ and */u/ were introduced into the second syllable, the former as "neutral" and the latter two as a harmonic pair, obviously first through certain sequential mergers in deverbal nouns

(Hakulinen 1968:40). Notice that deverbal nouns with "neutral" vowels in the stem took */u/ and not */ü/ as the derivational suffix. This indicates that the restriction initiated by the */iCa/, */eCa/ types, viz. that weak vowels may not occur after other than weak vowels in the first syllable, was strictly adhered to (words like *pesu* 'washing', *itku* 'crying' etc. are often cited as apparent violations of PVH, cf. Itkonen 1945:175-178, Karlsson 1974:123-124).

The next change in the system was the development of the /double/ */öö/ in the first syllable, and the arrival of */ö/ into the second. At this stage the harmonic relationships were as shown in Figure 4.

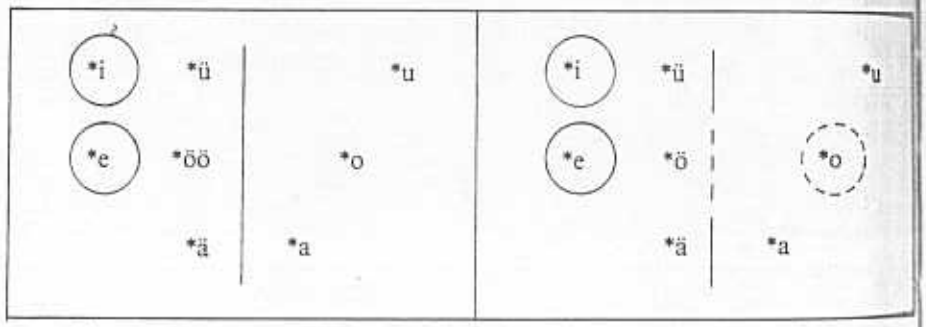

Figure 4. The vowel systems of the first and second syllables in approximately Proto-Finnish words. For further explanations see the text.

The dotted circle around the second syllable */o/ represents the beginning of the influence of the adjacent */ö/. The latter was obviously first introduced into the second syllable in words of the type */öCö/ (e.g. **tööhön* from earlier **tööhen*, illative sg. of **töö* 'work', cf. Hakulinen 1968:38), and since there were already words of the type weak vowel + * /o/ (e.g. **näko* 'sight'). * /ö/ (as against the neighbouring */o/) was not predictable after a weak vowel in the first syllable. What happened in the interest of accomplishing the predictability? In this case not only did the */äCo/ type become unproductive, but in the existing words of this type the second syllable */o/ changed to */ö/ (the number of such words that may have existed is not recoverable). Thus */o/ no longer appeared after initial weak vowels, and */ö/ attained predictability in accordance with Tendency (1), and as */o/ changed from unbounded to bounded, it at the same time changed from "neutral" to (*back*) harmonic. This is yet another proof of the claim that strong vowels are restricted by

PVH only through and because of the neighbouring weak vowel (and an indication that the distinction between bounded and unbounded strong vowels is a useful one). Let me again point out that words of the type */iCo/ and */eCo/ (e.g. *liitto* 'alliance', *teko* 'deed') remained so also after the introduction of */ö/. Had these also changed they would have violated the new PVH restrictions (cf. "neutral" vowel + */u/ combinations above).

The last change before essentially Modern Finnish was the arrival of /single/ */ö/ also into the first syllable. Interestingly, this seems to have taken place through the following developments: in a limited number of words of the type */eCü/, obviously coined in disrespect of the productive PVH restrictions (although another, more likely reason might be that the various developments did not take place simultaneously in different parts of the obviously geographically and culturally dispersed linguistic communities now referred to as single protolanguages), the */e/ of the first syllable changed to */ö/ (Hakulinen 1968:34-35). The change, while as such indicating that the word had been ill-structured from the beginning (or had been coined in dialects in which */e/ had not yet been "neutralized"; remember that the type */eCa/ which brought about the latter change was borrowed), also shows that the new restriction (weak vowels only after initial weak vowels) was becoming firmly established. The same is indicated by the behaviour of the handful of "neutral" vowel + /ü/ words still persisting in Modern Finnish (some of which are rather recent coinages). These, especially those with semantically opaque derivations, e.g. *kesy* 'tame' (the only native word of this type known to me for which no plausible derivational history has been suggested), *levy* 'board' and *city* tend to become /kösü/, /lövü/, and /sütü/, /süti/, /kitu/ etc. instead of /sitü/ (or the "correct" form /siti/) in the mouths of children and adults unfamiliar with any other language than Finnish (for an account of this tendency in terms of *markedness* see Wiik 1975:42-43; for a list of the relevant words (including /iCü/ words) see Groundstroem 1971:117-119). Having been firmly established in the language, the PVH restrictions now constitute inviolable (or hard-to-violate) "rules of pronounceability" for the speakers; for a recent summary of the many facets of Finnish PVH and a discussion of the classification of the PVH rules in terms of various parameters, see Campbell (1980).

Let me now summarize the co-occurrence restrictions of Modern Finnish vowels in acoustic terms. The combinatory possibilities of the three relevant classes of vowels are shown in Figure 5. The vow-

els within a class can be freely combined with each other. Between the classes, the arrows point from the first syllable vowels to those permitted in the second syllable, the broken arrow representing the existing, synchronically unproductive word types (»neutral« vowel + /ä/, /ü/) and inflected forms of stems containing only »neutral« vowels. The co-occurrence restrictions can be seen as the absence of arrows.

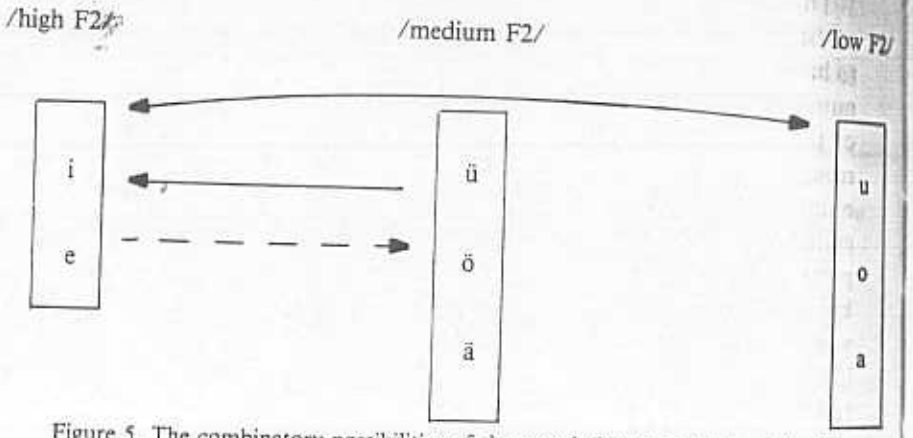

Figure 5. The combinatory possibilities of the vowel classes of Modern Finnish. For further explanations see the text.

The restrictions can be verbalized as follows. Taking all synchronically occurring combinations into account, it can be stated that

(a) /low F2/ vowels and /medium F2/ vowels cannot occur in the same word.

Excluding the synchronically unproductive combinations, the statement in a sense captures what the "language has been aiming at":

(b) A /low F2/ vowel cannot follow an initial /medium F2/ vowel, and a non-initial /medium F2/ vowel must be preceded by an initial /medium F2/ vowel.

Figure 5 and the above verbalizations permit several important conclusions. First of all, the harmonically relevant classes can be defined acoustically in terms of a single multivalued dimension F2 in comparison to the three dimensions *front-back, low-non-low* and *rounded-unrounded* necessary in an articulatory description (see e.g. Karlsson 1971:60-63). Secondly, also the co-occurrence restrictions themselves are all statable in terms of F2 alone; in (b) above this is possible because all of the unproductive types have been identically

treated. Thirdly, strong vowels can be freely combined with each other as predicted by the Tendencies. Fourthly, all restrictions concern weak vowels. More specifically, the situation described under (b) can be seen as the (as yet unattained) result of an historical development from the conflicting situation described under (a), and the developments clearly have as their goal the full predictability of non-initial weak vowels. Finally, the so-called "labial harmony" according to which /i/ and /e/ cannot be followed by /ü/ or /ö/ is only part of the more general restriction concerning all weak vowels following initial "neutral" vowels. Notice that the virtual lack of uninflected words violating "labial harmony" is not accidental in the present perceptual framework (cf. Itkonen 1945:178, Karlsson 1971:65-68, 1974:121, Wiik 1975:36-43).

To close this lengthy section, let me summarize some of the difficulties connected with the palatal assimilation (or "palatal attraction", see e.g. Kettunen 1960:44) theory of the causes and origin of PVH. According to this view the non-initial weak vowels (and hence PVH) have come about as a result of a progressive palatal assimilation, the *palatal* vowels of the first syllable having changed the once *back* vowels of the subsequent syllables to their modern *front* harmonic pairs. I will illustrate the difficulties involved in such a view by picking up /i/, the undeniably most *palatal* vowel of the lot, as the vowel in the first syllable. According to the palatal attraction theory the word type */iCä/ reconstructed for Proto-Uralic must have been ?/iCa/ at an earlier stage, changing later to */iCä/ through progressive palatal assimilation. Now the first inexplicable circumstance is this: how could ?/iCa/ exist at all for some period, why was it not "born" /iCä/ as it seems to defy the allegedly universal law of palatal attraction? Secondly, if the assimilatory force was for some time suppressed, then what caused the suppression? Thirdly, what may have ended the suppression? Fourthly, why did palatal attraction not affect the new */iCa/ type arriving after Proto-Uralic? Fifthly, how could this type become the productive combination instead of the older all-*palatal* */iCä/? Finally, why does the same supposedly universal, blind phonetic force fail to apply in so many languages of the world (in addition to PVH languages themselves)?

The origin of PVH in Uralic languages, i.e. the question of how the */å/-*/ä/ alternation of Proto-Uralic came into existence, is beyond the methods currently available for historical linguistics. However, it seems to have been established that */ü/ was introduced into the second syllable as the harmonic pair of */u/ at the outset,

the former developing (see Hakulinen 1968:40) in deverbal nouns with initial *front* harmonic vowels and the latter after *back* ones. Thus, */ü/ was not the result of earlier */u/ later changing to */ü/ after *palatal* vowels in the first syllable: */ü/ was there right from the beginning. As for */ö/, it seems to have entered the second syllable as a result of a process whereby */e/ following */h/ changed to the vowel of the first syllable in inflectional endings (e.g. **vööhön* from earlier **vööhen*, illative sg. of **vöö* 'belt', Modern Finnish *vyö*). This change can be described as an instance of progressive assimilation, to be sure, but it was certainly not caused by palatal attraction since the change took place in all kinds of words, i.e. */e/ also changed to e.g. */u/, */o/ and */a/ (Hakulinen 1968:38). Only after the way for */ö/ had thus been paved did e.g. the deverbal nouns of the type */äCo/ change to */äCö/ and other combinations of weak vowel + */ö/ become available. Finally, remember again that the alternants chosen after the *palatal* vowels */i/ and */e/ were the definitely non-*palatal* */u/ and */o/.

2. VOWEL HARMONY IN TURKISH

It is customary in phonological discussions to present the Turkish vowel system in terms of three articulatory dimensions as shown in Figure 6.

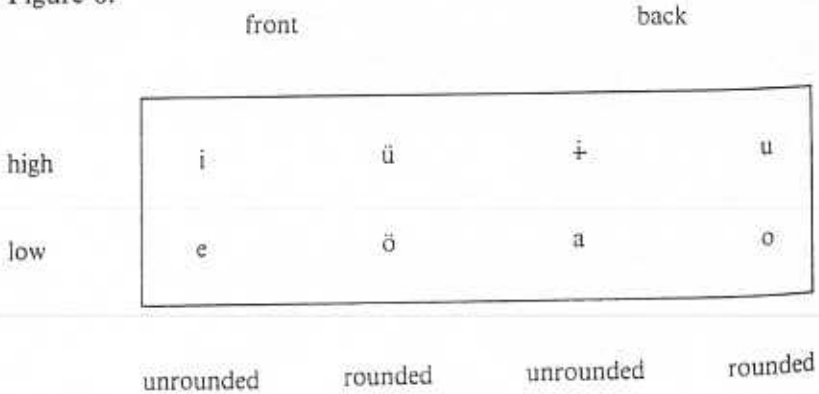

Figure 6. A traditional three-dimensional articulatory classification of the Turkish vowels.

The co-occurrence restrictions obtaining between the first and second syllables also hold between any subsequent syllables (/o/ and /ö/ occurring, however, only in the first syllable of the word). Thus, only the first two syllables need to be explicitly discussed here.

The standard view is that Turkish has two kinds of vowel harmony, viz. *palatal* harmony and *rounding* harmony. These are usually expressed by (factual equivalents of) the following rules:

(1) *Palatal* harmony: all vowels within a word must be either *front* or *back*.

(2) *Rounding* harmony:

(a) A *high* vowel must agree with the preceding vowel with regard to *rounding*.

(b) A *low* vowel in a non-initial syllable may not be *rounded*.

The combinations of vowels permitted by the above rules are visualized in Figure 7 in which the arrows again point from the vowel in the first syllable to that in the second and in which a square around a vowel indicates a homo-vocalic combination (i.e. /iCi/, /üCü/ etc.).

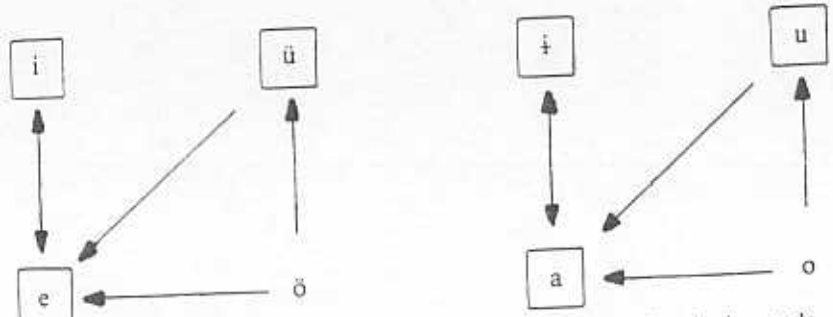

Figure 7. The combinatory possibilities of the Turkish vowels. For further explanations see the text.

The motivation for the standard phonological arrangement is immediately apparent in Figure 7: there is a perfect parallelism between the *front* and *back* sets of vowels with regard to the permitted combinations, and *palatal* harmony is clearly shown by the strict separation of the two sets. However, the account suffers from certain inadequacies as a descriptive statement. Thus, notice first that rules (1) and (2) are partly interdependent to the extent that (2a) is not a valid statement on its own since it would allow the combination of /i/ and /ɨ/ (both *unrounded*) and of /ü/ and /u/ (both *rounded*) in the same word. It is only the simultaneous imposition (or ordered application) of *both* (1) *and* (2a) that leads to the correct exclusion of the ungrammatical combinations. The dependence of *rounding* harmo-

ny on *palatal* harmony in this way already seriously undermines the claim that the two are separate phenomena. The claim is implicitly there, of course, through the use of different features to describe the phenomena, and it is made explicit by the adoption of two different names for the allegedly different kinds of harmony. Secondly, I am not certain whether (2b) (mentioned by e.g. Lewis 1967:15 and Crothers and Shibatani 1980:64) rightly belongs to a statement of the co-occurrence restrictions since it simply follows from the fact that /o/ and /ö/, the only *non-high rounded* vowels in Turkish, only occur in the first syllable. This is a restriction dependent on syllable position within the word and not on the vowel in the first syllable, and a phonotactic restriction of this sort should be stated somewhere else in the grammar.

Let me now proceed to an acoustically based characterization of the restrictions. Figure 8 shows a schematic reference arrangement of the vowels as they are grouped into natural classes. The approximate positions of the vowels have been determined from formant measurements by Sirén (1981) (for qualitative articulatory descriptions of Turkish vowels see Ruhlen 1975 and Crothers 1978 and the further references there). The enclosed area in Figure 8 contains the weak vowels.

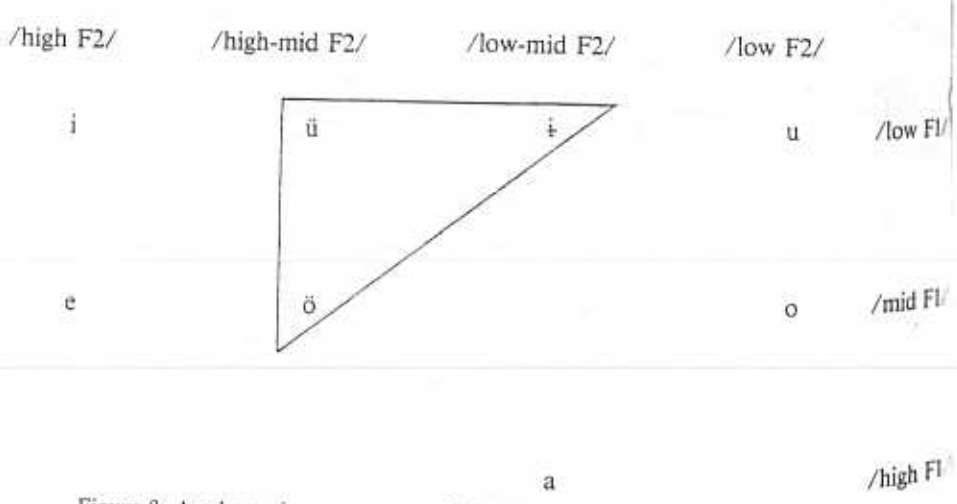

Figure 8. A schematic arrangement of the Turkish vowels in terms of F1 and F2. For further explanations see the text.

Using the acoustic specifications the restrictions can be stated as follows:

(3) A non-initial /low F1/ vowel must have the same /F2/ as the preceding vowel.

(4) A non-initial /non-low F1/ vowel must have either the same /F2/ as the preceding vowel or an /F2/ one degree higher.

These rules permit all and only the grammatical combinations also permitted by rules (1) and (2), assuming a prior statement to the effect that /e/ is the only /mid F2/ vowel capable of occurring in the non-initial syllables (which is not a vowel co-occurrence restriction). To see the practical effect of the restrictions more clearly, consider the following list of the permitted disyllabic word types, arranged so as to indicate which vowels may follow each initial vowel:

/iCi/	/üCü/	/ɨCɨ/	/uCu/
/iCe/	/üCe/	/ɨCa/	/uCa/
/eCi/	/öCü/		/oCu/
/eCe/	/öCe/		/oCa/
		/aCɨ/	
		/aCa/	

An examination of the list reveals that, as a result of the severe restrictions, the /F2/ of *all* non-initial vowels is completely predictable from the initial vowel and the /F1/ of the non-initial vowel itself. Thus while all permitted vowel distinctions after a given initial vowel involve a difference in /F1/, only some of them are accompanied by a concomitant /F2/ difference, and no non-initial vowel distinction depends on /F2/ alone.

In the descriptive framework using articulatory features, then, the restrictions are treated as belonging to the separate phenomena of *palatal* and *labial* harmony, statable in terms of three binary oppositions. Moreover, although explicit reference to two kinds of harmonies is made and different features are used in stating them, the two statements nevertheless turn out to be interdependent (cf. above). In an acoustically based description, in contrast, the restrictions emerge as a unitary phenomenon with a single result, an /F2/ harmony among the vowels in a word (as anticipated by Nearey 1980:239). The acoustic framework uses two multivalued dimensions (cf. the artificiality of describing differences in vowel *height* in terms of binary oppositions). Restrictions (3) and (4) are valid statements independent of each other; this, however, is not because two separate phenomena are involved but because the same /F2/ restriction al-

fects different sets of non-initial vowels in a different way. Notice further that (3) and (4), by restricting the occurrence of vowels in non-initial syllables only, correctly capture the "left-to-right" nature of the co-occurrence restrictions in an explicit manner.

The behaviour of the weak vowels and bounded strong vowels is in perfect agreement with our Tendencies (1) and (2). The predictions of Tendency (3) are not borne out as /e/ and /a/, unbounded by our definition, cannot follow all vowels in the initial syllable. While there are indications that the Tendency is not totally invalid for Turkish (see below), there may be room for another, qualifying tendency which is consistent with the total perceptual theory. Thus, assuming that small F2 differences between vowel phonemes in non-initial syllables are not desirable on perceptual grounds, it is not hard to accept that as the F2 dimension becomes more crowded, stronger predictions concerning the identity of such vowels from the first syllable become necessary. This in mind, I propose the following additional tendency which at least seems to account for the difference between the behaviour of the unbounded strong vowels of Finnish and Turkish:

> Tendency (4): The greater the number of non-initial weak vowels on the F2 dimension in a language, the stricter the dependencies of the non-initial vowels on those of the initial syllable.

Thus, the fact that the F2 dimension is divided into four linguistically relevant degrees in Turkish compared with the three in Finnish may explain why the restrictions are altogether more severe in the former (four degrees also seems to be the universal upper limit in any position, see Crothers 1978:117-121). As far as the *high* vowels of Turkish are concerned, this looks straightforward enough, but there are no perceptual grounds for restricting /e/ and /a/ (both being sole occupants of their respective F1 levels in the non-initial syllables). Perhaps the fact that the latter belong to harmonic classes with the *high* vowels is as such sufficient to account for the restrictions, i.e. there may be a general principle according to which the same restrictions apply to all members of an harmonic class irrespective of their individual characteristics.

Further evidence of the relative independence of strong vowels in comparison to weak ones comes from certain marginal sections of Turkish vocabulary (the following information is taken from Lewis 1967:17). Firstly, there are about a dozen simple native words which

contain both *front* and *back* vowels. The *back* vowel is invariably /a/ and the *front* vowel is either /e/ or /i/. These are all strong vowels, and no such combinations involving weak vowels are mentioned by Lewis. Secondly, there is a group of invariable suffixes, about half a dozen, all containing strong vowels exclusively. Thirdly, words of foreign origin with only strong vowels seem to retain their approximate original vowel qualities although they violate the Turkish co-occurrence restrictions, whereas those with weak vowels in the original language in non-Turkish combinations tend to change so as to conform to the native pattern. In sum, to the extent that minor exceptions to the main restrictions exist, the marginal words tend to contain only strong vowels whereas exceptional words containing weak vowels are not tolerated.

3. A PERCEPTUAL THEORY OF THE CAUSES OF PVH

The explanation offered below must be considered tentative at least for three reasons. Firstly, the two-formant acoustic description of vowel quality adopted is only an approximation of the perceptual properties. Secondly, the data basis of the paper is limited. Thirdly, the theory involves an amount of speculation not supported by experimental findings. Observe the obvious methodological difficulties facing any attempts to test the proposed theory directly on speakers of PVH languages: the aim is to find out what has caused the phenomenon, but speaker-hearers have internalized the predictabilities involved and this directs their perception. Similarly, speakers of any language have of course acquired the sound pattern of their native tongue and this influences their perception. This is not to say that the theory is untestable: the claims made are empirical issues subject to revision or rejection as new data are gathered. On the whole, existing accounts of PVH (of which I have of course only seen a small section) appear as overly formal to my temperament, and one of the motivations for writing this paper was the wish to stimulate more substance-oriented approaches.

It was seen in the preceding sections that the working Tendencies are highly consistent with the data. Implicit in the Tendencies is the claim that F1 is perceptually more salient than F2. The perceptual asymmetry receives corroboration from several sources. Consider first the very distinction between strong and weak vowels. The typological fact that the former predominate in vowel systems (for quan-

titative data, see Crothers 1978) is readily explainable on the basis of the claimed asymmetry and the assumption (?) that languages tend to first make use of such vowel sounds as are most clearly distinct from the other vowels: strong vowel systems are frequent because they involve only large F2 differences. Secondly, Flanagan's investigations (1955) of just noticeable differences of vowel quality indicate that, on the average, a change of about 75 Hz is needed in F2 to get an 80 per cent change of judgments of vowel quality, the change in F1 needed to obtain a similar shift of judgments being about 30 Hz, i.e. less than half as counted in Hz. Thirdly, investigations of perceptual confusions among vowels show that misperceptions tend to result from incorrect estimations of the F2 of the stimulus vowels, vowels with approximately the same F1 being most often confused with each other (Lafon 1968, Suen and Beddoes 1972). Fourthly, algorithmic quantitative models designed to generate the empirically established typological dispersion of vowels in the acoustic F1-F2 space tend to produce too many vowels whose mutual differences are based on F2 alone, i.e. the number of weak or interior vowels is clearly greater than that attested in natural languages. Thus the model of Linjencrants and Lindblom (1972), using a mel scale and assuming that the perceptual contributions of F1 and F2 as measured in mels are equal, generated too many weak vowels situated between [i] and [u] because, as thus measured, this dimension seems to have plenty of room for intervening vowels. The model has been later revised by Lindblom (1975; unfortunately, I have not had access to this unpublished paper) and by Crothers (1978) who, after reducing the F2 dimension to about half the scale it had in the acoustic space as used by Linjencrants and Lindblom, succeeded in reducing the discrepancy between the empirical dispersion of vowels and theoretical predictions. This was obviously done more and less *ad hoc* (Crothers 1978:126-127), but it is interesting to note that the conversion of linear Hz values to the perceptually more accurate Bark scale (Zwicker and Feldtkeller 1967:92-94) has the same effect as the reduction adopted by Crothers: the weight of F2 relative to F1 is reduced by a factor of about two, the range of total F2 version then being about twice the total variation of F1 as against a fourfold difference when estimated in terms of linear Hz values. Even so, F1 still remains about twice as strong perceptually as F2. Recently, at least a partial resolution of the [i] — [u] problem has been demonstrated by Lindblom who, making use of certain psycho-acoustically motivated scales (although these mainly derive from investigations using

pure tone signals as stimuli), transformed acoustic representations of steady-state vowels into corresponding representations in the auditory periphery and thereby reduced in a non-*ad hoc* manner the distance between the two vowels (Lindblom forthcoming, see also Bladon and Lindblom 1981).

Another claim implicit in the Tendencies is that the first syllables of words are easily locatable in the flow of speech by the hearer for use in the prediction of the vowels in the subsequent syllables. This in turn presupposes the easy locatability of word boundaries. As speech is used for communicating meanings and as words are important meaning-bearing elements, the locatability of word boundaries is in fact a precondition for successful communication. The problem is rather to find out *how* this is accomplished. At least the majority of PVH languages so far reported have a fixed word stress (the Uralic languages mostly stressing the first and the Altaic languages the last syllable; for minor exceptions consult e.g. Ruhlen 1975), and it is customary to say that such an automatic, non-phonemic stress functions as a word boundary signal. Hence it might be tempting to conclude that word boundaries are signalled by word stresses in such languages. However, such a conclusion is hardly warranted since, as is well known, the prosodic cues of word stresses are often absent from connected speech (word stress in a sense being the "place" within a word in which an eventual syntactically, pragmatically etc. determined sentence stress is realized). Notice also that it is well-nigh impossible to divide speech in an unknown language into word-size chunks. This could be due to differences in the cues used to signal word boundaries in various languages, but it may also indicate that, on the whole, word boundaries are not systematically realized phonetically in connected speech. The recognition of words from fluent speech has been recently investigated by Cole and Jakimik (1980), who conclude that their research "provides at least partial support for each of the following statements:

> Words are recognized through the interaction of sound and knowledge.
> The words in an utterance are recognized one after another.
> A word's recognition locates the sounds which begin the following word.
> The sounds in the beginning of a word are used to access word candidates.
> The sounds in a word are processed sequentially.
> A word is consciously recognized when the sequential analysis of

sound eliminates all word candidates but one." (Cole and Jakimik 1980:161).

The above conclusions have been reached for English and it remains to be assessed to what extent similar principles hold for word recognition in other, typologically different languages. At the same time, one can hardly deny the intuitive plausibility of such a view of word recognition.

The fact that the prosodic cues potentially signalling word stress are often blurred by higher-order prosodic influences in fluent speech has important repercussions on PVH. In languages with a moving or phonemic stress the correct location of the stressed syllable is of course important, and in face of the unreliability of the strictly prosodic cues these languages must resort to other, segmental cues. Among these are restrictions on the choice of vowels permitted in the unstressed and stressed syllables. I suggest that the functional motivation for such restrictions (which seem to be extremely common in languages with a moving stress) is the ensuing contribution to the locatability of word stress by the hearer, and it is only natural that the perceptually less salient weak vowels are among those first prohibited in the unstressed syllables (in which, to better enhance perceptual prominence of the stressed syllable, the vowels are often extensively reduced in quality). In languages with a fixed, fully predictable word stress, on the other hand, there is no semantic motivation for similar restrictions. Thus, instead of asking why particular vowels occurring in the obviously most permissive first syllable also occur in the non-initial syllables, it might be more appropriate to ask why not. In the absence of the demands of moving word stress there is no obvious reason, and this may well account for the spread of weak vowels into the non-initial syllables in PVH languages. In other words, it is suggested that it is a general characteristic of languages with non-phonemic word stress that all vowels occurring in the language tend to spread to all syllable positions, an empirical question difficult to decide at present because of the lack of sufficiently specific descriptions of such languages. Sammallahti (1980:6) considers the lack of balance between the vowel systems of the first syllable and the second syllable in Proto-Uralic to be one of the factors which caused the whole system to develop into the nearly perfect balance in Modern Finnish (the second syllable lacks some of the diphthongs occurring in the first). In the light of the foregoing I readily agree with Sammallahti, adding only that in a sense the development can be regarded as a syntagmatic extension of the tendency towards

the maximization of the information content of the speech signal (i.e. the more syllables in the word with the maximum paradigmatic system, the greater the number of potential different meanings conveyed per syntagmatic unit). However, when weak vowels are introduced into the non-initial syllables, perceptually motivated PVH restrictions enter the scene.

Theoretically, languages can be divided into four groups on the basis of whether they permit weak vowels in the non-initial syllables of the word and whether they have PVH restrictions, as shown below.

	weak vowels in non-initial syllables	*PVH*
(a)	no	no
(b)	no	yes
(c)	yes	no
(d)	yes	yes

The evaluation of the number of languages belonging to each group is again hampered by the fact that, even for languages for which phonological/phonetic descriptions are available, these are often not specific enough about the phonotactic restrictions. Nevertheless, group (a) can safely be said to include the majority of the languages of the world (Ruhlen 1975:317-319, Crothers 1978:106-108) as weak vowels in any syllable position only occur in the minority of languages. As for group (b), notice that it is perfectly possible to imagine a language without weak vowels but with PVH, e.g. one with the five universally most common vowels (Crothers 1978:104 and *passim*) in which the co-occurrence restrictions were as shown below, the vertical line again indicating the division into harmonic sets.

first syllable		second syllable	
i	u	i	u
e	o	e	o
a		a	

This elegant system, with *front* and *back* vowels constituting mutually exclusive sets, would thus be most easily statable in terms of the current articulatory features, and such systems would be highly expected if the avoidance of vowels occupying opposite positions on the *front-back* dimension indeed were the cause of PVH. To my knowledge, however, no languages with such a system (PVH but only strong vowels in non-initial syllables) have been reported. This is

another empirical question but my claim, based on the perceptual theory, is that a language with such a set of non-initial vowels would have no motivation for PVH restrictions.

There are languages belonging to our group (c). Notice, however, that I am not claiming that the presence of weak vowels in non-initial syllables necessarily causes PVH. The claim is, instead, that PVH does not develop without weak vowels in these positions. The latter is a weaker claim, impying that weak vowels can be tolerated in non-initial syllables by making them perceptually sufficiently salient (so as to be easily distinguishable from the neighbouring strong vowels) also by means other than PVH (cf. the discussion of moving word stress languages above: some of the Germanic languages would be a case in point). In other words, I wish to maintain that the presence of non-initial weak vowels is a necessary but not necessarily a sufficient precondition to the development of PVH. As such, then, the existence of languages belonging to group (c) does not constitute counter-evidence against the present framework (although the mechanisms by which non-PVH languages manage to solve the problem of non-initial weak vowels are relevant also to PVH: further research in this area is clearly needed).

As far as I can judge, all PVH languages belong to group (d). The class of PVH languages currently known includes only languages belonging to the Uralic and Altaic families and the languages could perhaps be enumerated: what I am trying to do is to specify some of the phonetic conditions for and against the development of PVH in the belief that PVH is not arbitrary in this respect but, rather, the lawful product of general substantive constraints on the shape of possible sound patterns. In other words, I wish to contribute to the "attempt to derive linguistic form as a consequence of various substance-based principles pertaining to the use of spoken language and its biological, sociological, and communicative aspects" (Liljencrants and Lindblom 1972:859).

We are now in a position to at least tentatively answer the question made above as to why PVH is so rare among the languages of the world, and the answer is simply that the circumstances provoking PVH seldom coexist in the same language. Given that both weak vowels and fixed word stress are necessary prerequisites for the development of PVH, the number of potential PVH languages is immediately drastically reduced when only the former precondition is considered in the light of typological information. Of the 693 languages contained in Ruhlen's (1975) sample, 80 (or about 12 per cent) are reported to have *front rounded* vowels and 66 (about 10 per cent) *back unrounded* vowels (in many cases the same languages are included in both counts). Without going into a detailed language by language analysis of Ruhlen's data here, and concentrating on the reportedly more common *front rounded* weak vowels only, it may be observed that languages belonging to the Altaic and Uralic families (i.e. those mostly having also fixed word stress and PVH) account for no less than 48 per cent (38 languages) out of the 80 languages reported to have *front rounded* vowels in any syllable position. This proportion strongly suggests the dependence of PVH on weak vowels, especially when taken together with the apparent non-existence of PVH languages with only strong vowels. It remains for further detailed investigations to assess to what extent the absence (?) of PVH in the remaining languages with weak vowels can be explained with reference to the word stress pattern.

The perceptual theory of the causes of PVH, then, can be summarized as follows. The immediate reason for the co-occurrence restrictions within a word known (somewhat misleadingly as the above analyses show) as palatal vowel harmony is the presence of so-called weak vowels in the non-initial syllables of the word. The scope of PVH is the word because the word is the basic meaning-bearing unit of language. The introduction of weak vowels into the non-initial syllables is caused by the need to increase the potential meaning content (lexical efficiency) of the speech signal. The tendency to spread all vowels occurring in the most permissive position to all positions within the word is, it was tentatively suggested, common to all languages with a fixed word stress. However, since weak vowels are inherently less salient perceptually than the strong vowels typically adjacent to them, the spreading of weak vowels into all syllable positions (in those fixed stress languages that have weak vowels at all) would, as such, necessitate an amount of F2 computing per word which is not easily tolerated because of the relatively lesser perceptual efficacy of F2 differences as compared with those of F1. Therefore, the occurrence of weak vowels in non-initial syllables has to be restricted so as to make their identification easier: obviously only so many accurate estimations (perhaps only one estimation?) of F2 per word can be effectively performed by the listener under normal circumstances (here I on purpose ignore the multitude of non-phonetic factors operating in normal communicative interaction). Notice that a listener in a communicative situation, on receiving incoming speech, tries to come up with a match between the acoustic signal

and words stored in his mental lexicon: what I am doing is to specify the limits on the possible form of such lexical entries, or phonetic conditions for the acceptance of stretches of speech into the vocabulary. To resume, the restrictions make the identification of weak vowels predictable from the vowel in the first syllable and the perceptually more efficient acoustic dimension of F1. The vowel in the first syllable, indicated to the listener by the completion of the preceding word and (occasionally) by a clear manifestation of the fixed word stress (or sentence stress falling on that syllable), is then paid special attention to by the listener. On the one hand, the listener examines the initial vowel to specify its identity and, more specifially, to determine its /F2/ correctly, thus activating the less sensitive F2 detection mechanism in this particular position (which is pointed towards, as needing particular attention, by higher-order semantical recognition processes). Notice that in moving-stress languages the syllable demanding extra attention would be pointed to by partly different indicators, i.e. the phonetic cues of the semantically most important syllable would be different. On the other hand, the /F2/ information of the first vowel is used by the listener (of a PVH language) as a basis for forming preliminary hypotheses concerning the likely /F2/ of the subsequent vowels, initial hypotheses rendering some candidates improbable. If the identification of the first vowel with regard to its /F2/ is fully conclusive, then the perceptual mechanism is freed from paying detailed attention to F2 in the subsequent syllables. It is very likely that the information received prior to the recognition of the whole word is retained in short-term memory, and it is not necessary that the identity of the initial vowel has to be fixed before that of the subsequent ones. This is obviously true in any language, and particularly in PVH languages since the division of vowels into mutually exclusive sets also enables an amount of backward prediction (e.g. in the event of noise during the first vowel). In the normal case, however, the recognition of the vowel in the first syllable presumably precedes that of the later incoming ones in real time. In the non-initial syllables, then, the perceptual mechanism usually has to pay attention to approximate F2 values only (in addition to F1). The mechanism for deriving accurate F2 estimates is re-activated on the arrival of a message (from wherever lexically, syntactically, pragmatically etc. determined recognition decisions are made) informing that the current word has been identified (which, see Cole and Jakimik 1980, often happens before the speaker has reached the end of the word). To the extent that the phonetic and non-phonetic sources of

information in the receiving end are in conflict, the communicative intention of the speaker may fail or the listener may detect a break of PVH.

In conclusion, notice that the only difference in word recognition between speakers of PVH languages and other languages postulated by the theory is that the former use predictions concerning the /F2/ of non-initial syllables as an aid whereas the latter do not, and that this directly relates to a structural (i.e. linguistic) difference between the sound patterns of these languages. Contrast this to the speaker differences necessarily presupposed by an articulatory explanation of the causes of PVH: speakers of PVH languages are for same reason affected by the need to avoid the combination of *front* and *back* vowels (and this in a particularly selective manner) while speakers of other languages are obviously not affected!

ACKNOWLEDGEMENTS

My thanks go to Olli Aaltonen for useful discussions of vowel perception (which went far beyond the crude outlines dealt with in the paper); to Kaisa Häkkinen for her generous and expert advice on all matters concerning the history of the Finnish vowels — while she is in no way responsible for the errors that still remain, her clarifications have greatly reduced their number; and to an anonymous referee in the editorial board for numerous and welcome suggestions for improvements to an earlier draft.

REFERENCES

Anderson, Stephen R. 1980: Problems and Perspectives in the Description of Vowel Harmony. In Vago, Robert M. (ed.) 1980: *Issues in Vowel Harmony*. John Benjamins, Amsterdam, pp. 1-48.

Bladon, R.A.W. and Lindblom, Björn 1981: Modeling the Judgment of Vowel Quality Differences. *Journal of the Acoustical Society of America* 69, 1414-1422.

Campbell, Lyle 1980: The Psychological and Sociological Reality of Finnish Vowel Harmony. In Vago, Robert M. (ed.) 1980: *Issues in Vowel Harmony*. John Benjamins, Amsterdam, pp. 245-270.

Carlson, R., Granström, B. and Fant, G. 1970: Some Studies Concerning Perception of Isolated Vowels. *Speech Transmission Laboratory — Quarterly Progress and Status Report (KTH, Stockholm)* 2-3/1970, 19-40.

Cole, Ronald and Jakimik, Jola 1980: A Model of Speech Perception. In Cole, Ronald (ed.) 1980: *Perception and Production of Fluent Speech*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 133-163.

Crothers, John 1978: Typology and Universals of Vowel Systems. In Greenberg, Joseph H. (ed.) 1978: *Universals of Human Language. Volume 2. Phonology*. Stanford University Press, Stanford, California, pp. 93-152.

Crothers, John and Shibatani, Masayoshi 1980: Issues in the Description of Turkish Vowel Harmony. In Vago, Robert M. (ed.) 1980: *Issues in Vowel Harmony*. John Benjamins, Amsterdam, pp. 63-88.

Fant, Gunnar 1973: *Speech Sounds and Features*. Mouton, The Hague.

Flanagan, James L. 1955: A Difference Limen for Vowel Formant Frequency. *Journal of the Acoustical Society of America* 27, 613-617.

Groundstroem, Axel 1971: Studien zur phonologischen Komponente der modernen finnischen Hochsprache. Doctoral dissertation, University of Stockholm.

Hakulinen, Lauri 1968: *Suomen kielen rakenne ja kehitys*. Third edition, Otava, Helsinki.

Häkkinen, Kaisa 1981: Die Stellung des genuinen Wortschatzes in der finnischen Sprache der Gegenwart. *Finnish-Ugrische Mitteilungen* 5, 177-199.

Itkonen, Erkki 1945: Onko kantasuomessa ollut keskivokaaleja? *Virittäjä* 49, 158-182.

Itkonen, Erkki 1948: Vokaalikombinaatiot ja vartalotyypit. *Virittäjä* 52, 124-144.

Janhunen, Juha 1981: Uralilaisen kantakielen sanastosta. *Journal de la Société Finno-Ougrienne* 77, 219-274.

Karlsson, Fred 1971: *Finskans rotmorfemstruktur: en generativ beskrivning*. Publications of the Phonetics Department of the University of Turku 10, Turku.

Karlsson, Fred 1974: *Centrala problem i finskans böjningsmorfologi, morfofonematik och fonologi*. Suomi 117:2, Suomalaisen Kirjallisuuden Seura, Helsinki.

Kettunen, Lauri 1960: *Suomen lähisukukielten luonteenomaiset piirteet*. Mémoires de la Société Finno-Ougrienne 119, Helsinki.

Lafon, Jean-Claude 1968: Auditory Basis of Phonetics. In Malmberg, Bertil (ed.) 1968: *Manual of Phonetics*. North-Holland, Amsterdam, pp. 76-104.

Lewis, G.L. 1967: *Turkish Grammar*. Oxford University Press, Oxford.

Liljencrants, Johan and Lindblom, Björn 1972: Numerical Simulation of Vowel Quality Systems: the Role of Perceptual Contrast. *Language* 48, 839-862.

Lindblom, Björn 1972: Phonetics and the Description of Language. In Rigault, André and Charbonneau, René (eds.) 1972: *Proceedings of the Seventh International Congress of Phonetic Sciences*. Mouton, The Hague.

Lindblom, Björn 1975: Experiments in Sound Structure. Paper read at the Eighth International Congress of Phonetic Sciences, Leeds.

Lindblom, Björn 1980: Economy of Speech Gestures. In MacNeilage, Peter (ed.) 1980: *Speech Production*. Springer, Heidelberg.

Lindblom, Björn (forthcoming): Phonetic Universals in Vowel Systems. In Ohala, John (ed.): *Experimental Phonology*. Academic Press, New York.

Lindblom, Björn and Sundberg, Johan 1971: Acoustical Consequences of Lip, Tongue, Jaw and Larynx Movement. *Journal of the Acoustical Society of America* 50, 1166-1179.

Linell, Per 1979: *Psychological Reality in Phonology*. Cambridge University Press, Cambridge.

Nearey, Terrance M. 1980: On the Physical Interpretation of Vowel Quality: Cinefluorographic and Acoustic Evidence. *Journal of Phonetics* 8, 213-241.

Ruhlen, Merrit 1975: *A Guide to the Languages of the World*. Language Universals Project, Stanford University.

Sammallahti, Pekka 1979: Über die Laut- und Morphemstruktur der Uralischen Grundsprache. *Finnish-Ugrische Forschungen* XLIII, 22-66.

Sammallahti, Pekka 1980: Introduction to a symposium on vowel harmony held at the Fifth International Finno-Ugric Congress in Turku 1980. In Ikola, Osmo (ed.) 1980: *Congressus Quintus Internationalis Fenno-Ugristarum, Turku* 20.-27. VIII. 1980, *Pars* III. Suomen Kielen Seura, Turku, pp. 3-7.

Sirén, Kirsti 1981: Difficulties in the Perception of English Vowels: A Contrastive/Error Analysis of English and Turkish. Unpublished Master's thesis, English Department, University of Oulu.

Suen, Ching Yee and Beddoes, Michael P. 1972: Discrimination of Vowel Sounds of Very Short Duration. *Perception and Psychophysics* 11, 417-419.

Trubetzkoy, N.S. 1971: *Grundzüge der Phonologie*. Fifth impression. Vandenhoeck and Ruprecht, Göttingen.

Wiik, Kalevi 1965: *Finnish and English Vowels*. Annales Universitatis Turkuensis B 94, Turku.

Wiik, Kalevi 1975: *Vokaalisoinnun ongelmia*. Publications of the Phonetics Department of the University of Turku 14, Turku.

Zwicker, Eberhardt and Feldtkeller, Richard 1967: *Das Ohr als Nachrichtenempfänger*. Hirzel, Stuttgart.