A Formal Theory of Vowel Harmony

J.-R. Vergnaud

A FORMAL THEORY OF VOWEL HARMONY

J.-R. Vergnaud

0. In this paper, I shall discuss the theory of vowel harmony that is developed in On the form and interpretation of phonological rules (J.R. Vergnaud, forthcoming) and in Formal phonology (Halle, Prince, and Vergnaud, forthcoming). This theory was sketched in a paper that I gave at the Fifth International Congress of Logic, Methodology and Philosophy of Science in August 1975 (see Formal properties of phonological rules, forthcoming). First, I shall survey the notions and formal devices that are proposed and elaborated in these papers.

1. The papers just mentioned take as their point of departure the theory of phonological rules outlined in Chapter 8 of The Sound Pattern of English (SPE; for ease of exposition, I shall refer to this theory as the standard phonological theory) and they discuss various inadequacies of this theory; they, then, present and justify a revision of the standard theory. Specifically, the modifications have to do with the formalization of phonological statements that refer to long components (in the sense of Z. Harris) and/or to discontinuous elements: a phonological model is put forth which is nonlinear (in the sense of N. Clements, 1976: Vowel harmony in nonlinear generative phonology; a phonological representation is linear if it can be exhaustively analyzed into an ordered sequence of units having no ordered subparts; a model is linear iff. its phonological representations are linear). Here, I shall discuss briefly some of the considerations that are developed in the aforementioned papers.

The formal devices that are part of the theory of language must meet several conditions of adequacy. A first requirement, as we know, is that they permit us to formulate general statements about the language that are true and significant, and that they provide a basis for distinguishing these from other generalizations which are false, or which are true but not significant. A second requirement is that they permit us to write a theory of grammar that is explanatorily adequate. I shall discuss the

University of Massachusetts Occasional Papers in Linguistics, Vol. 6 [1980], Art. 7

latter requirement, in connection with the theoretical proposals put forth below, at the end of this paper. In this first part of my presentation, I will try to show that there are simple considerations that lead one to question the descriptive adequacy of the standard theory and to consider nonlinear alternatives to the standard theory. That is, the standard theory does not meet the first requirement above, but there exists a nonlinear model, namely the one I shall present below, that meets it.

Consider vowel harmony in Nez Perce. I shall follow here the description in SPE. The words of Nez Perce fall into two classes with regard to their utilization of vowels; in the words of the first class the vowels are selected from the set [i a o]; in the words of the second class the vowels are chosen from the set [i æ u]. The morphemes themselves constitute two mutually exclusive categories: morphemes of the first category appear in words of the first class only, whereas morphemes of the second category appear in words of both classes. Hence, morphemes of the first category show no vowel alternations and select their vowels from the set [i a o], whereas morphemes of the second category exhibit the vowel alternations a-æ and o-u, depending on whether the morpheme appears in a word of the first or the second class. Let us call [+H] the set of vowels [i a o] and [-H] the set [i æ u]. The facts just sketched can be accounted for by assuming that the vowels in the underlying representation of a morpheme of the first category are [+H], and that the vowels in the underlying representation of a morpheme of the second category are [-H], and by postulating a rule (or a well-formedness condition) that distributes the feature [+H] to all vowels of a word containing at least one [+H] vowel. This phonological rule (or condition) appears to be a very natural one. In particular, similar harmony systems are found in many Bantu languages and in such languages as Diola-Fogny, Somali, Kalenjin, etc.; these are what Aoki (1968) calls "asymmetric" vowel harmony systems. Note that it is very easy to state the phonological rule (or condition) above in English: the description of the rule we gave was simple, natural, and perspicuous. It is easy to see that this simplicity, this naturalness, and this perspicuity cannot be matched by the notation of a linear model, such as the standard theory for example. Within the latter theory, we have to posit the following rule:

$$(1)\quad V \longrightarrow [+H] \;/\; \left\{ \begin{array}{l} \#X\ [+H]\ Y \underline{\qquad} \\ \underline{\qquad} Z\ [+H]\ W\# \end{array} \right\}$$

Specifically, the harmony system of Nez Perce has two properties that make it difficult to describe within a linear theory: first, in each word, it concerns a subset of the set of segments of the word which is not a continuous substring of the word, namely the set of vowels of the word; second, it is a bidirectional process. I shall consider these two properties in turn. First, consider the fact that (1) affects a discontinuous substring of the word. What the standard theory (or every linear model for that matter) lacks is a way of representing such discontinuous substrings. A natural way of representing such strings is to use the formalism of the trees. Figure 1 gives an idea of what we have in mind. In Figure 1, $\#Q_1V_1Q_2 \ldots C_iV_i \ldots Q_n\#$ where V_i, $i=1,\ldots,n-1$ is a vowel and Q_i, $i=1,\ldots,n$ is a (maximal)

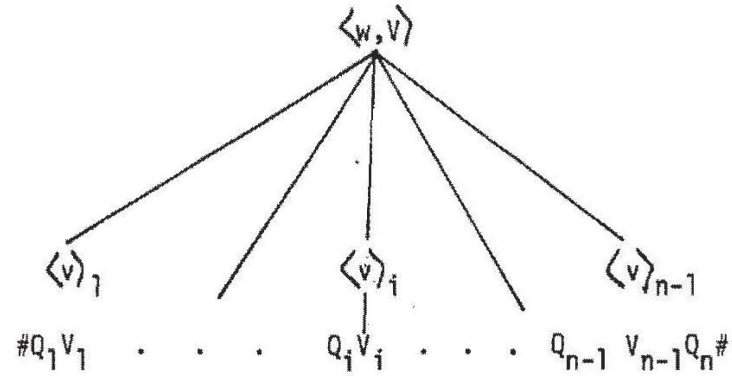

Figure 1

string of non-syllabic segments, is the standard representation of a word. Each V_i is dominated by a "node" $\langle V \rangle_i$ ("V" in "$\langle V \rangle$" is the feature [+syllabic]. The sequence $\langle V \rangle_1 \langle V \rangle_2 \ldots \langle V \rangle_{n-1}$ is dominated by the "node" $\langle w, V \rangle$ ("w" here stands for "word"; I will come back to this below). In some sense, each $\langle V \rangle_i$ is a "preterminal node"; the "terminal" string dominated by $\langle V \rangle_i$ is the standard representation of the vowel V_i (that is, it is a unit or an archi-unit). We see that it becomes possible, e.g., to write a phonological statement that refers to the set of vowels of a word: with respect to this formalism, the set of vowels of a word is a constituent, namely $\langle w, V \rangle$. Which solves the first problem that arose in connection with the formalization of the harmony system of Nez Perce. Consider now the second problem, namely the bidirectionality of the process. I will show that the formalism I have just discussed permits us to solve this problem elegantly. Suppose that phonological nodes can acquire the feature specifications by means of phonological rules or conventions; for example, the node $\langle w, V \rangle$ in Figure 1 might acquire the feature specification [+H] (such a specification should be distinguished from the feature content of the node, which defines the category of the node - in the latter example, this feature content is [+syllabic] -; a feature specification assigned by a rule (or convention) could be viewed as an indexing of the node by the specified feature under consideration). Suppose furthermore that there exists a universal convention that stipulates that a unit which is dominated by a node "indexed" by [αF] must be [αF] (this convention could be a well-formedness condition or a rewriting rule; I shall discuss this in greater detail below). Then, the harmony system of Nez Perce can be described by the following statement:

(2) $\langle w, V \rangle$ has the index [+H] iff it dominates a [+H] vowel.

It is easy to show there exists a simple way of expressing (2) within a theory that incorporates the formalism of trees (in particular, we can write a rule which involves none of the complexities displayed by (1)). We will delay till later the precise formalization of Nez Perce vowel harmony. Informally, the phenomenon can be described as the indexing of a node by a (specified) feature. A node that can be so indexed, I shall call an *harmonizing node*. A feature that can be an index on a node, I shall call an *harmonic feature*. For example, in Nez Perce, $\langle w,V \rangle$ is an harmonizing node and [±H] is an harmonic feature. The phenomenon under consideration, namely the indexing of an harmonizing node by an harmonic feature, I shall call *α-harmony*. The rest of this paper will be devoted to a discussion of α-harmony. There is another kind of harmony, which I call β-harmony. This second type of harmony is found in, e.g., Takelma, Southern Agaw, Tshiluba. For reasons of time and space, I will not discuss β-harmony here. I refer the reader to my forthcoming paper and to *Formal Phonology*.

2. α-harmony can be further analyzed into various subtypes. A parameter in the definition of these subtypes is the form and the content of the harmonizing node. Up to this point, we have only discussed the node $\langle w,V \rangle$. Of course, our theory defines other types of nodes. I shall now make the formalism a little more explicit. First, let us analyze the meaning of the symbol "$\langle w,V \rangle$". I said earlier that "w" meant "word"; furthermore "V" stands for "[+syllabic]". The definition of "$\langle w,V \rangle$" is as follows:

(3) a maximal word-internal sequence of nodes $\langle V \rangle$ such that each $\langle V \rangle$ in the sequence dominates V *is a* $\langle w,V \rangle$ where V=[+syllabic]

For the sake of this discussion, I shall assume that every unit that belongs to N, where N is some natural class, is dominated by a node $\langle N \rangle$. Of course, such structures are most easily representable in a three-dimensional space. But, this is not relevant to our discussion. We can see that (3) is only a special case of a more general definition which involves three parameters:

(4) a maximal D-internal sequence of nodes $\langle N_1 \rangle$ such that each node $\langle N_1 \rangle$ in the sequence dominates a unit belonging to N_2, where N_1 and N_2 are natural classes and N_1 includes N_2, *is a*...

(3) corresponds to the values D=w, N_1=N_2=V. Suppose that we take D in (4) to be w and that we make N_1 and N_2 vary, with the constraint N_1=N_2. Then, we obtain various types of nodes which may be relevant to the kind of processes we are discussing. To take an example, consider the Navajo rule of strident assimilation, described by Z. Harris in *Structural Linguistics* (1960). In a Navajo word, strident obstruents agree in anteriority. The rule of assimilation is a regressive one (it is an optional rule, whose application is conditioned by such factors as the rapidity of speech). Thus, we have Figure 2. The harmonizing node is $\langle w,N \rangle$, where N is the natural class [+obst, +strid, +cor] and the harmonic feature is [±ant]. The specification of the feature [±ant] that indexes the harmonizing node is the one dominated by the leftmost node $\langle N \rangle$ under $\langle w,N \rangle$: this is what the arrow in Figure 2 means. Note that this harmony process is formally different from the one found in Nez Perce. In particular, the relevant occurrence of the harmonic feature in Navaho (that is, the occurrence of the harmonic feature that indexes the harmonizing node) is determined by the structural

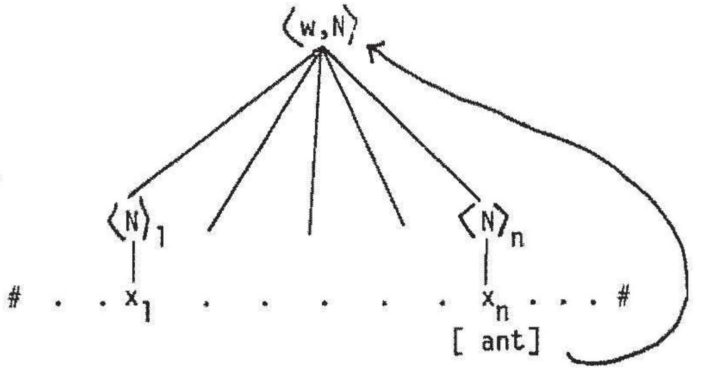

Figure 2

description of a rule, which is not the case for Nez Perce. I shall come back to this difference later on.

To illustrate further the definition in (4), consider the case where D=w and where N_1 = [+segment]. The, we can define a node noted $\langle$w, N_2, [+seg]$\rangle$: $\langle$w,N_2, [+seg]$\rangle$ dominates a continuous word-internal string of segments that are N_2; specifically, $\langle$w,N_2, [+seg]$\rangle$ dominates a maximal string of this type. Consider for example the fast speech rule of nasal spread found in a form of Castilian Spanish spoken in Southern Spain (this example is drawn from N. Clements, to appear). In this language, such a form as <u>una revista</u> [una ɹeβiʰta] may become [ũnã ɹ̃ẽβ̃ĩʰta] in fast speech. In other words, all the units in a maximal continuous string of sonorants that contains a nasal segment are nasalized. This is typically a case of α-harmony. The harmonizing node is $\langle$w,N_2, [+seg]$\rangle$, where N_2 is [+seg, +son], and the harmonic feature is [±nasal]. Up to this point, we have been essentially discussing certain aspects of the theory of phonological representations I have in mind. In the next section I shall turn to a discussion of the theory of vowel harmony proper. Of course, many more things could, and should, be said about phonological representations. In particular, it is clear that there will be substantive constraints that will delimit the class of possible nodes, and formal constraints that will govern the relations between the nodes in a phonological representation. On this, see the articles mentioned at the beginning of this paper. Note that the trees we have been discussing are very different from the trees found in Syntax. In particular, their hierarchical organization is quite shallow (there is some amount of it, however, as we shall see).

3. I shall now describe the formal conventions that I propose to account for vowel harmony. For the sake of illustration and reasoning I shall use an abstract example which could be viewed as an abstract version of, let us say, Akan. This abstract vowel harmony system, I shall call L. Let [F] be the harmonic feature in L. I am assuming that vowel harmony in L is centrifugal, determining the value of affix vowels from the root outward. For the sake of this discussion, I shall assume that L does not have neutral vowels. However, L has opaque vowels. The latter notion will be defined formally below.

I shall introduce the system of formal conventions in an axiomatic fashion and I will show that it accounts for L. Then, I will show that this system can be extended so as to cover the asymmetric vowel harmony systems (such as Nez Perce). We shall see that, in all these cases, no rule of grammar is involved: the behaviour of these vowel harmony systems is governed by universal conventions. To conclude this section, I shall discuss briefly the case of rule-governed harmony (such as Navajo).

Consider then L. I shall assume a phonological representation slightly more complex than the one posited earlier for Nez Perce (cf. Figure 1). Specifically, I assume that the "vowel tree" correspo to a word is as in Figure 3: P corresponds to the prefixes, S, to the suffixes, and R, to the root. P and S are left-branching and right-branching, respectively; they are both "binary-branching" (thus, if an affix has two vowels, its vowels are not dominated by sister-preterminal nodes). A and B are the first vowel and the last vowel of the root, respectively. I shall assume that each element of the phonological representation, be it a node or a unit, bears an index which is its "name"; these indices or names are integers. I shall assume furthermore that the harmonic feature [F] is ternary: it can be specified +,-, or 0. I argue in Vergnaud (forthcoming) that this position is theoretically consistent (see also N. Clements, 1976). In what follows, the notation "x_i" represents the specification of F in the element i of the phonological representation. Let K be a non-opaque vowel in an affix, our theory assumes that the value of x_k in underlying representation is 0_k; in other words, a non-root non-opaque vowel is unspecified for F in underlying representation. We define an harmonizing node to be a node that is indexed by [0F]; this indexing is a lexical property of the nodes; the range of possibilities is determined by Universal Grammar. In particular, the root-node R may or may not be harmonizing; the first case corresponds to the case of "regular" roots. A root-vowel or an opaque vowel is specified + or - in underlying representation. Our theory contains the following universal "linking" conventions:

(5) (i) LC I

$0_i \longrightarrow \alpha_i$, $\alpha_i = \alpha_j$, where j is a unit and i immediately dominates a node that immediately dominates j.

(ii) LC II

$\alpha_k \longrightarrow 0_k$, where k is a unit which is immediately dominated by a node q such that q = + or -

To illustrate, suppose that the harmonizing suffix node S_i immediately dominates a node $\langle V \rangle_j$ that dominates a unit V_j such that α_j = + or -; then, S_i gets indexed [+F] or [-F] (accordingly) by LC I and the specification for F in V_j is replaced by 0 (by LC II). Of course, we are assuming that LC I precedes LC II. LC I and LC II apply to the underlying representation and after every application of a phonological rule. In some sense, the system LC I, LC II constitutes an "exchange" system that converts

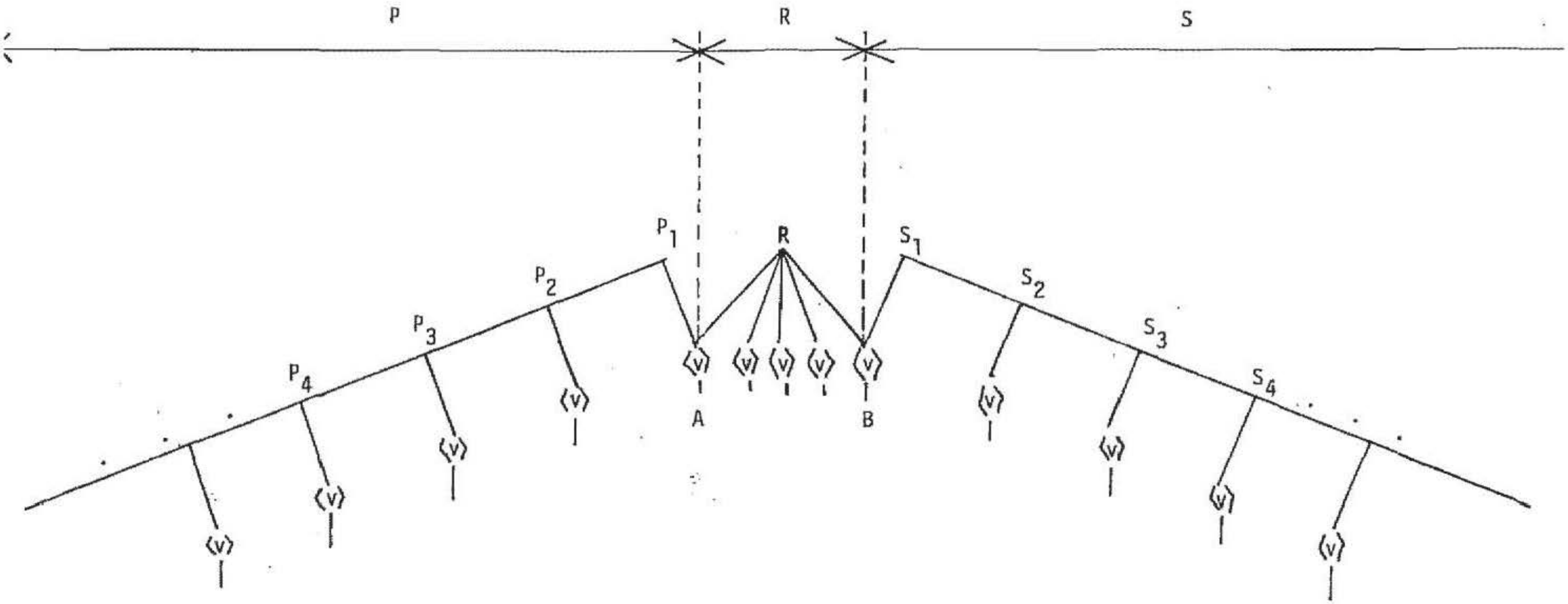

Figure 3

141

a standard segmental representation into a "prosodic" representation or an "autosegmental" representation. I shall assume that [OF] can be rewritten as [+F] or as [-F] only by rules of grammar or be convention LC I. Then, the output of the phonological component may contain occurrences of F that are unspecified (i.e. specified o). If there were no other convention, the output of the phonological rules then would be ill-formed, because it could not be interpreted by the phonetic rules. I shall posit the following convention:

(6) C III

Let j be a unit that is immediately dominated by a node that is immediately dominated by k, or let it be a node that is immediately dominated by k, then O_j is interpreted as equal to α_k by the phonetic rules. We write: $O_j \equiv_P \alpha_k$.

It is easy to show that the model described above accounts for L. Consider now the asymmetric vowel harmony systems, e.g. Nez Perce. The difference between Nez Perce and L is that in Nez Perce the occurrences of the feature [+H] that are opaque are not bound by lexical morphemes; that is, every occurrence of [+H] in every morpheme is "opaque", or, so to speak, the feature [+H] is opaque <u>en-soi</u>. The difference between Nez Perce and L then is to be described in terms of the properties of the "linking" conventions. Specifically, LC I does not apply to Nez Perce, but is replaced by LC III:

(7) LC III

$O_i \longrightarrow \alpha_i, \quad \alpha_i = \alpha_j$, where i immediately dominates j

I am assuming that LC III can apply iteratively.

Finally, consider strident assimilation in Navajo. In this case the indexing of the harmonizing node is effected by a phonological rule and not by convention LC I. Everything else remains the same.

4. To conclude, I have presented a new formalism that permits us to write a very restrictive theory of vowel harmony. I have surveyed the basic formal mechanisms of the latter theory. This theory of vowel harmony makes it possible to distinguish clearly what is universal from what is language particular in harmony systems. For example, we see that a child that is learning the harmony system of Nez Perce has only three things to learn: first, that the harmonizing node is w,V ; second, that the harmonic feature is [H]; third, that [+H] is opaque. Note that our theory permits us to separate the formal properties of harmony systems from their substantive properties, which is a considerable advantage. In some sense, our theory could be viewed as a formalization of the theory of vowel harmony presented in N. Clements (1976). I believe that the theory I have discussed in this paper entails the most significant properties of Clements' model, and makes precise some of its mechanisms. Note, in particular, that any precise autosegmental account of harmony would have to include conventions similar to LC I and LC II.

A last remark. We see that the notation we have discussed at the beginning of this paper is adequate not because it restricts significantly the class of possible grammars, but because it permits us to formulate a theory that restricts the latter class, which is typically the situation one would expect in a mature field.

Bibliography

Aoki, H. 1968. Toward a Typology of Vowel Harmony, IJAL 34. 142-145.
Chomsky, N., and M. Halle, 1968. The Sound Pattern of English. New York: Harper and Row.
Clements, G.N. 1976. Vowel Harmony in Nonlinear Generative Phonology.
Clements, G.N. To appear. The Autosegmental Treatment of Vowel Harmony.
Halle, M. and J.-R. Vergnaud. Forthcoming. Metrical Structure in Phonology. Unpublished ms., M.I.T.
Harris, Z. 1960. Structural Linguistics. The University of Chicago Press.
Kean, M.-L. 1975. Markedness in phonology. Ph.D. Thesis, M.I.T.
Vergnaud, J.R. To appear. Formal Properties of Phonological Rules, in Proceedings of the 5th International Congress of Logic, Methodology, and Philosophy of Science, 547-566. eds. R.E. Butts and J. Hintikka.
__________. Forthcoming. On the Form and Interpretation of Phonological Rules.