

NOTES ON THE DEVELOPMENT OF PHONOLOGICAL THEORY

Stephen R. Anderson, Dept. of Linguistics, U.C.L.A., Los Angeles, California, U.S.A.

In describing the state of phonological theory in recent years, Basbøll distinguishes between "substance based" and "formal" approaches to the fundamental problems of definition and explanation in the field. This is undoubtedly a useful opposition, and one that corresponds to most people's intuition about what is at issue in some recent controversies. My own work has been primarily in the direction Basbøll would characterize as "formal", and I would like therefore to describe the issues involved from that perspective. I am sure my fellow co-reporter will do justice to the other side.

I would agree with Basbøll that most phonologizing in recent years has been carried out within a comparatively unitary set of assumptions about the defining problems of the field, and therefore that a single broadly construed school of phonology has dominated research (despite efforts to promote comparatively minor differences of opinion to the status of fundamental differences). Whether acknowledged or not, most of the problems dealt with in this school are set (or at least foreshadowed) in the 'standard theory' of Chomsky and Halle's Sound Pattern of English (SPE). I have given elsewhere an account of recent developments in connection with the details of that program (cf. Anderson, 1979), and will not repeat that discussion here. I will instead confine myself to some remarks of a more general nature.

To gain perspective on the issues involved in recent phonological debate, it seems to me quite useful to consider the parallels between the evolution of phonological theory and that of the study of the foundations of mathematics. Let us recall that the primary nature of a phonological theory, as expressed in SPE, is the development of an explicit formal notation for phonological description. In combination with an evaluation function for grammars defined over this notation, this would result in a comprehensive axiomatization of the subject matter of phonology, in the sense that all problems connected with the discovery of a correct (or 'descriptively adequate') account of sound structure in a given language would thereby be reduced to the mechanical manipulation of expressions in a fully explicit notational system. Of

course, SPE does not claim to have accomplished this goal, but it is nonetheless the program of the theory. The successes achieved within this framework were seen as confirmation of the plausibility of such an axiomatization.

The program of SPE is thus strikingly similar to that of another fundamental work of 20th century thought, Whitehead and Russell's Principia Mathematica (PM). That work developed a program of reducing all of the intellectual content of mathematics to the formal manipulation of expressions in a logistic system by means of fully explicit rules. While the calculus of formal logic in which PM proposed to express mathematical propositions is of course quite unlike the descriptive apparatus for phonological expressions envisaged by SPE, the goal of expressing all of the content of a field in terms subject to formal manipulation by well-established rules is common to the two works.

PM's account of the foundations of mathematics was initially greeted enthusiastically, since it promised to give a full reconstruction of the traditional notion that the truth of mathematical propositions derives from logic alone, and not from contingent facts about the world. This enthusiasm rapidly gave way to dissatisfaction, however, as it became apparent that there were fundamental obstacles to the logicist program. In particular, the theory in its basic form was seen to give rise to a number of the paradoxes which had long troubled mathematicians (such as various forms of the problem of the barber who shaves everyone who does not shave himself, and others). In order to remedy this difficulty, Russell had proposed what is known as the theory of 'types', roughly speaking a restriction on the kinds of classes that can be referred to in a given expression. Unfortunately, the theory of types itself had the undesirable consequence of rendering unstatable or meaningless many basic propositions in number theory. It was thus necessary, in the full system of the PM, to appeal to an axiom of infinity and an axiom of reducibility, whose plausibility and intuitive appeal are vastly less than that of the rest of the logical system. Since the theory of types seemed unavoidable in the context of the logic of the PM, and since it seemed to lead to such counterintuitive emendations of the system, the logicist program for the foundations of mathematics was gradually abandoned.

Partially in response to the perceived failure of this approach, other views of the foundations of mathematics were developed on other assumptions. Among the most important of these alternative views was that presented by Brouwer and others under the title of intuitionism. A primary tenet of this school is the rejection of all expressions purporting to refer to objects that cannot in fact be fully constructed. In particular, expressions that refer to explicitly infinite sets are disallowed, since (while one can give directions for indefinitely enlarging the extension of a set) it is obviously not possible to complete the enumeration of such an object. This has the immediate consequence that the fundamental paradoxes that arise for Russell's system are avoided, since the problematic classes turn out to be impossible to construct within the limits of an intuitionist logic.

Intuitionists have attempted to reconstruct as much as possible of the subject matter of mathematics while adhering to such limitations. In many cases, it turns out to be possible to reformulate classical results in such a way as to be able to derive them in these terms. In other areas, however, this is impossible, and the intuitionists are then led to conclude that such areas of mathematics are in fact meaningless: a somewhat controversial result.

In the course of developing the intuitionist program, its practitioners have clearly revealed much about the conceptual basis of mathematical propositions. This program does not really lead to independent advances, however, since it provides the basis for only a partial development of mathematics. Relatively few working mathematicians seem willing to accept the limitations on their subject matter imposed by the premises of intuitionist logic, and thus although they can be said to have shed light on a (proper) subset of the field, the intuitionists cannot be said to have replaced the traditional modes of inference for mathematics as a whole.

A similar development can be traced in phonology. In particular the program of SPE leads, in the end, to the result that considerations of the substantive phonetic content of representations and rules has no natural role in the system of phonology. This problem is recognized in the famous chapter 9 of SPE, where a solution is proposed in the form of the theory of markedness.

Such a theory is in fact an attempt to reduce exhaustively the considerations of phonetic content that might be relevant to phonology to purely formal expression in the notation. While it too was greeted with much initial enthusiasm, it is noteworthy that essentially no substantial analyses of phonological phenomena have appeared subsequently in which this aspect of the theory plays a fundamental role. This seems to be due at least in part to the fact that the set of 'marking conventions' required to account for the facts of one language or group of languages simply do not extend to comparable utility in others. The purely mechanical problems encountered here are immediately apparent to anyone attempting to formulate a description in such a way, and as a result serious efforts to take account of phonetic content have generally been pursued along quite different lines.

If we would draw the full lesson from these observations, it seems to me that we must conclude that the role of phonetic content in phonology is such as to reveal a fundamental inadequacy in the full 'logician' program for the field sketched in SPE. The theory of markedness, that is, seems to be an emendation of the same character as Russell's theory of types. The lesson in each case is not that a consistent formal system of the required character cannot be constructed, but rather that the only available ways of doing so inevitably lead to fundamental conflicts with the subject matter which the theories are intended to account for. Neither a logical basis for mathematics nor a comprehensive notation for the expression and comparison of phonological descriptions are proven to be wrong: they are simply shown to be incomplete in essential aspects as full reconstructions of the domains of thought with which they are concerned.

In reaction to the inadequacies of the account of phonetic substance offered by SPE, a similar 'intuitionist' approach (though not really in the form of a coherent school) has grown up in phonology, in attempts to remedy the presumed paradoxes resulting from the standard theory by restricting its conceptual richness. Most notably, the approach of Natural Generative Phonology (NGP) has been to require the reconstruction of phonological accounts without appeal to abstract entities or to putatively counterintuitive logistic principles such as relevant explicit ordering. This constitutes a retreat from idealism to a theory

founded insofar as possible on what are (from a linguist's point of view, if not that of an experimental psychologist) the observable and immediately verifiable aspects of linguistic structure. As such, it is immediately reminiscent of the constructivist basis of intuitionist mathematics.

In fact, the parallel is quite close. NGP succeeds in reconstructing a large part of the traditional domain of phonological description, though sometimes in unfamiliar terms. In doing so, it has shown us much about the conceptual basis of more familiar solutions. On the other hand, there are also many aspects of what has usually been taken to be phonology which are inaccessible on its premises. These areas of phonology are either written off altogether (that is, declared to be linguistically meaningless) or ascribed to the operation of vague, nonphonological principles (such as 'via-rules', essentially a name for the description of those aspects of phonology that cannot be accounted for without an appeal to abstract entities).

Now a consistent adherent of NGP may well be happy with the result that certain domains are thereby eliminated from consideration, just as a confirmed intuitionist may be convinced of the result that much of classical and modern mathematics is literally meaningless, but in both areas traditional, pre-systematic practitioners of these subjects have felt discontent with the portion of their fields that can be treated within such a radically 'constructivist' account. If NGP must, as argued in critiques such as that of Gussmann (1978), throw out the baby with the bath water, many phonologists would resist the contention that a priori considerations of psychological reality make this way of avoiding the disregard of phonetic substance characteristic of SPE the right line.

Now in mathematics, the disillusionment with the full logicist program which followed from certain aspects of the system of PM certainly did not have the result that serious work in formal mathematical logic came to a halt. On the contrary, the sort of investigation carried out in these terms turned out to constitute an interesting and coherent field of study, defining significant problems of its own to which solutions could be sought that would result in essential contributions to our understanding of the structure of mathematics. If it is not possible to decide all

mathematical questions within this field, it is still an area of basic importance, concerned with very real problems.

It seems to me that the situation in phonology is entirely analogous. The formalist program of SPE is undoubtedly incomplete as the basis of a comprehensive account of all problems in phonological structure in natural language. It still appears to constitute a well-formed and important subpart of that study, with real problems in its own right that can be formulated, addressed, and decided, and which lead to basic improvements in our understanding of the nature of sound systems. It is in this area, indeed, that I think we are still (largely due to the monumental results represented by SPE) best equipped to make substantial progress. Our growing awareness of the range of problems that cannot be reduced to notational decisions, in fact, has the effect of refining our understanding of the contribution made by those results that can be obtained. In this respect, my own (admittedly quite partisan) evaluation is that the advances that can be made by taking formal questions seriously far exceeds the interesting but limited scope of reductionist efforts such as that of NGP.

As an example of such a question, let us briefly consider the problem of whether or not morpholexical ('word formation') processes necessarily precede purely phonological processes in grammars. It should be emphasized that the notion "precedes" in this formulation of the issue is not a purely metaphoric (or metaphysical) one, nor is its validity dependent on an interpretation in terms of temporal sequential processing, either in speakers' production or in historical change. Rather, it refers to the issue of whether or not there are ever morpholexical processes whose operation crucially depends on (and thus presupposes the presence of) information about a form which is only supplied by the generalizations represented by some phonological process - and which is thus unavailable in the underlying representations of forms. The device of sequential application is a particular formalization of this, but it should be kept in mind that it is the relation of informational dependency that is at issue.

The value of this observation for our knowledge of language, however, turns on the fact that it is logically a contingent proposition. Simply asserted by fiat, it becomes totally uninteresting, a limitation on what sort of world we are willing to countenance.

Taken otherwise, however, it can be falsified by the demonstration that in at least one language there is a well-supported instance of a contrary dependency. Such examples are not, in fact, especially difficult to document. A particularly interesting (because highly structured) case is found in Javanese (cf. Dudas, 1974). In this language, the 'elative' (a sort of intensive form) of adjectives is constructed by replacing the last vowel of the word by a tense high vowel: i if the basic vowel was front and non-round, u if the basic vowel was back and round. Thus we find alternations such as luwe 'hungry', elative luwi; adoh 'far', elative aduh, and many others. If the final vowel is a, however, there are two cases: if the last syllable is closed, the elative is formed in i, as in gampang 'easy', elative gamping. If the last syllable is open, however, the elative is formed in u. Thus, from underlying /kamba/ 'insipid', the elative is kambu. The explanation of this difference is not far to seek, however. A general phonological rule of the language neutralizes the opposition between /a/ and /o/ in final open syllables, replacing both by o. This rule is responsible for alternations such as dino 'day', dinane 'the day' (from the root /dina/), and is dependent only on phonological information for its operation. There is much more to be said about these rules, and about others with which they interact, but I think those who consult Dudas' paper and the sources to which she refers will find that this account does not distort the situation. Now in fact the behavior of basic /a/ in elative formations is clear: it is precisely where this vowel would be replaced by o (in final open syllables) that elative formation treats it in the same way as back rounded vowels (like /o/). Otherwise, it behaves like the other unrounded vowels. The generalization that is apparent in these data is that elative formation depends on the information that is supplied by the rule replacing final /a/ in open syllables by o, not on the underlying form directly. In other words, this rule of word-formation follows the phonological rule in question (as well as some others, as Dudas documents). Notice that this demonstration proceeds quite otherwise than by "considering the ... notation as given ... [and] drawing conclusions ... from the notation" as Basbøll seems to suggest. Rather, it is precisely the appropriate form of the notation that is at issue: in particular, an aspect of the organization of grammars concerning

mathematical questions within this field, it is still an area of basic importance, concerned with very real problems.

It seems to me that the situation in phonology is entirely analogous. The formalist program of SPE is undoubtedly incomplete as the basis of a comprehensive account of all problems in phonological structure in natural language. It still appears to constitute a well-formed and important subpart of that study, with real problems in its own right that can be formulated, addressed, and decided, and which lead to basic improvements in our understanding of the nature of sound systems. It is in this area, indeed, that I think we are still (largely due to the monumental results represented by SPE) best equipped to make substantial progress. Our growing awareness of the range of problems that cannot be reduced to notational decisions, in fact, has the effect of refining our understanding of the contribution made by those results that can be obtained. In this respect, my own (admittedly quite partisan) evaluation is that the advances that can be made by taking formal questions seriously far exceeds the interesting but limited scope of reductionist efforts such as that of NGP.

As an example of such a question, let us briefly consider the problem of whether or not morpholexical ('word formation') processes necessarily precede purely phonological processes in grammars. It should be emphasized that the notion "precedes" in this formulation of the issue is not a purely metaphoric (or metaphysical) one, nor is its validity dependent on an interpretation in terms of temporal sequential processing, either in speakers' production or in historical change. Rather, it refers to the issue of whether or not there are ever morpholexical processes whose operation crucially depends on (and thus presupposes the presence of) information about a form which is only supplied by the generalizations represented by some phonological process - and which is thus unavailable in the underlying representations of forms. The device of sequential application is a particular formalization of this, but it should be kept in mind that it is the relation of informational dependency that is at issue.

The value of this observation for our knowledge of language, however, turns on the fact that it is logically a contingent proposition. Simply asserted by fiat, it becomes totally uninteresting, a limitation on what sort of world we are willing to countenance.

Taken otherwise, however, it can be falsified by the demonstration that in at least one language there is a well-supported instance of a contrary dependency. Such examples are not, in fact, especially difficult to document. A particularly interesting (because highly structured) case is found in Javanese (cf. Dudas, 1974). In this language, the 'elative' (a sort of intensive form) of adjectives is constructed by replacing the last vowel of the word by a tense high vowel: i if the basic vowel was front and non-round, u if the basic vowel was back and round. Thus we find alternations such as luwe 'hungry', elative luwi; adoh 'far', elative aduh, and many others. If the final vowel is a, however, there are two cases: if the last syllable is closed, the elative is formed in i, as in gampang 'easy', elative gamping. If the last syllable is open, however, the elative is formed in u. Thus, from underlying /kɛmba/ 'insipid', the elative is kɛmbu. The explanation of this difference is not far to seek, however. A general phonological rule of the language neutralizes the opposition between /a/ and /o/ in final open syllables, replacing both by o. This rule is responsible for alternations such as dino 'day', dinane 'the day' (from the root /dina/), and is dependent only on phonological information for its operation. There is much more to be said about these rules, and about others with which they interact, but I think those who consult Dudas' paper and the sources to which she refers will find that this account does not distort the situation. Now in fact the behavior of basic /a/ in elative formations is clear: it is precisely where this vowel would be replaced by o (in final open syllables) that elative formation treats it in the same way as back rounded vowels (like /o/). Otherwise, it behaves like the other unrounded vowels. The generalization that is apparent in these data is that elative formation depends on the information that is supplied by the rule replacing final /a/ in open syllables by o, not on the underlying form directly. In other words, this rule of word-formation follows the phonological rule in question (as well as some others, as Dudas documents). Notice that this demonstration proceeds quite otherwise than by "considering the ... notation as given ... [and] drawing conclusions ... from the notation" as Basbøll seems to suggest. Rather, it is precisely the appropriate form of the notation that is at issue: in particular, an aspect of the organization of grammars concerning

the formalization of possible interdependences between rules. In this case, the answer seems clear. The proposed constraint is not a valid one, and must be replaced by some other, less restrictive (and hence, less interesting) one. The relation between such questions of formalism and the data of actual languages is quite direct, as is the contribution their resolution can make to our understanding of the organization of sound systems.

In contrast to this situation, however, the problem of how phonetic substance is related to formal description will only receive a serious answer when we recognize the possibility of a radical difference between them. In particular, the requirement that in order to have merit, a theory must be explanatory in the sense of being rigidly predictive imposes in essence the requirement that all questions of substance be expressible ultimately in a formal calculus manipulated by mechanical rules of inference. This sort of program, typified by the theory of markedness, has gotten more and more vague of late, but the requirement of predictability amounts to the demand that substance be reduced to a form commensurate with other, formalizable constituents of a phonological description.

It seems to me that this sense of predictability is inappropriate. The existence of distinct linguistic systems developed from a common antecedent through the differential operation of historical change, taken seriously, provides a falsification of its premises nearly as fundamental as Gödel's demonstration that there are propositions formulable within arithmetic whose truth value cannot be decided in principle within that system. Appeals to social factors and the like are at present mere hand-waving: the conviction that somewhere an explanation exists that will preserve predictability. We must recognize that it is precisely the character of phonetic substance to be both non-random and non-deterministic: a 'logical' formalization of its role in phonology is unavailable in principle. I have suggested elsewhere an alternative sort of goal, the attainment of an ex post facto understanding of phonological processes (or 'exegetic adequacy'), which is (at least, at present) more appropriate for phonology than the program of complete predictability.

When the principles of a theory lead to a domain of conflict, as for instance in the case of the Neogrammarian notions of Laut-

gesetz and Analogie, we certainly do not have predictability - but that does not mean we have not advanced our knowledge. We may well claim to understand the facts to a greater degree than we would in the absence of principles, despite the fact that we cannot claim that the facts could not have been otherwise. An excellent example of this situation is furnished by the current state of research into apparently well-motivated but mutually inconsistent principles that govern rule orderings in phonology.

The atmosphere of 'science', toward which we all aspire, tends to force us into a rather radical mechanism. This is useful when it makes us examine the conceptual bases of our work and to seek the regular connections among phenomena; but it may ultimately become sterile if we insist that only a completely deterministic account is worthy of consideration as 'scientific'. After all, if physics and mathematics can accept fundamental principles of indeterminacy, phonologists should be willing to countenance the uncertain as well.

Basbøll is surely right that an understanding of the role of substance in phonology can only come from an appreciation of the science of that substance, to wit, phonetics. Equally clear, however, not all of the results of phonetic research are equally applicable. It is an axiom of applied mathematics (though not, it sometimes appears, of all phoneticians) that "the purpose of computing is insight, not numbers"; and the most central sort of phonetic research is undoubtedly that which aims at a notion of phonetic motivation and explanation. The work of scholars such as Sweet, Passy, Grammont, and others of an earlier generation has somewhat fallen out of favor as unscientific, largely because of its non-deterministic character (though also on account of the charge of vagueness).

The most promising sort of synthesis seems to me to be found in the work of Baudouin de Courtenay, the 50th anniversary of whose death we mark this year. Baudouin's inspired integration of the explanatory role of traditional phonetics (in accounting for the entrance of low-level processes into the system) with that of the study of the internal structure of grammars (in treating the relations, both evolutionary and synchronic, among the various sorts of rules) deserves serious reconsideration (cf. deChene and Anderson, 1979). Such a synthesis is also one of the merits of Stampe and Donegan's 'Natural Phonology'.

The kind of understanding we can hope to achieve from phonological research, then, is arguably possible only if we abandon the ultimately unreachable goal of complete predictability. Attempts to achieve predictability by imposing arbitrary limits on the form of phonological descriptions, such as by the decision a priori that 'extrinsic' or language particular orderings are one kind of complexity that languages absolutely cannot tolerate, seem unmotivated and misguided. In the absence of an understanding of general cognitive processes underlying language that could explain them, such 'constraints' cannot be taken seriously as the motivation for particular decisions about the appropriateness of descriptions. As Basbøll notes, such extralinguistic explanation seldom plays a real (rather than rhetorical) role in phonological theorizing.

To me, however, this suggests that much of the actual research Basbøll characterizes as 'substance based' is ultimately unproductive, since it is based on the arbitrary imposition of restrictive principles which rule out otherwise well-motivated descriptions. We have no way of knowing a priori what sorts of complexity, abstractness, etc. are tolerated by natural languages, and the only way of discovering this is through the unbiased examination of the facts they present. This is not to deny that such programs can lead to significant insights, as in the case of their emphasis on a distinction between morphological and purely phonological rules, which has evidently led to major improvements in our understanding of sound systems. Nonetheless, far from suggesting that the study of formal problems in phonology, of the sort arising in the framework of SPE, should be abandoned, the lesson of this research seems to be that it is only by taking these formal matters seriously that their ultimate role in a comprehensive view of sound structure can be appreciated.

References

- Anderson, Stephen R. (1979): "On the subsequent development of the 'standard theory' in phonology", in Current approaches to phonological theory, D. Dinnsen (ed.), Bloomington, Indiana: Indiana University Press.
- deChene, Brent, and Stephen Anderson (1979): "On compensatory lengthening", Lg. 55.
- Dudas, Karen (1974): "A case of functional phonological opacity: Javanese relative formation", in Studies in the linguistic sciences 4:2, 91-111.
- Gussmann, Edmund (1978): Explorations in abstract phonology, Lublin: Uniwersytet Marii Curie-Skłodowskiej.

FORMAL AND SUBSTANTIVE APPROACHES TO PHONOLOGY

Joan B. Hooper, State University of New York at Buffalo, USA

The main report on phonological theory by Hans Basbøll gives a rather thorough treatment of current phonological research. This response will mention a few additional works and issues, but it is intended primarily to supplement Basbøll's report by presenting in somewhat greater depth an examination of the theoretical diversity underlying current phonological research. Our point of departure is the distinction Basbøll discusses between a "substance based" versus a "formal" approach to phonological research. This distinction characterizes quite broadly two major research trends in generative phonology, but leaves out some important differences. In order to highlight several theoretical positions, the "substance" versus "formal" distinction will be divided into two separate distinctions which cross-classify fully. This brief report will discuss the resulting categories and the type of research emanating from each of them. Basbøll noted that his classification of two types of phonology was only rough and ignored some individual differences. Similarly, the distinctions I will make are also rough, and are meant only as a useful organization of a diversity of research perspectives.

1. Two major issues

1.1. The most direct interpretation of the substance-formal distinction divides phonological research into that which investigates formal or structural properties of grammars and that which investigates substantive properties. The former research is concerned with levels of representation, and how they relate to one another, and with the formal properties of rules, and the formal relations among them. Substantive properties can be thought of roughly as content properties -- phonological features are the content of representations, and changes in phonological features in the presence of other related phonological features are the content of rules. For most investigators, the substance of phonology is phonetic (but see section 3.2.).

This aspect of the substance/formal distinction is not so much a theoretical issue as a distinction between two types of interests, which are not mutually exclusive. Most researchers would agree that phonology has both a formal and substantive side, and that the two need to be studied together at least to some extent.