On the Evaluation Measure

Paul Kiparsky
M.I.T.

The evaluation measure can be construed as an explication of the concept "learnability of the abstract system of a language" (Koefoed 1974). It interacts with the long-established factors of ease of perception and ease of production in setting the direction of linguistic change. It is evident that the three factors can impose mutually conflicting requirements on the grammar. For example, phonological processes, as well as syntactic transformations, serve to facilitate speech production and perception, but both generally do so at the cost of making the structure of the linguistic system harder to discover. Moreover, the functions of production and perception can themselves partially conflict with each other.

Natural languages appear to stay close to an optimum beyond which ease of production, perception, and acquisition can not be jointly increased, and improvements in any one respect entail sacrifices in the others. There is, however, no fixed point of balance between the three factors. In the first place, variation in the relative weight of the factors, depending on the demands imposed by the speech situation, creates a stratification of speech into functionally determined styles. Rapid speech forces maximization of speakability, and familiar speech also allows it by virtue of decreasing the conflicting perceptibility requirement. Formal speech, and speech over a noisy channel (e.g. a bad phone connection) require maximal perceptibility. There is also considerable variation in structural complexity, cf. the maximally simplified styles of speech used with children and foreigners (Ferguson 1971).

Secondly, alongside this kind of functionally governed stylistic variation there is fluctuation over time. The teleology of change is "local" in the sense that changes can be motivated by any single factor, and they are not, as far as has been shown, blocked if they have bad effects with respect to some other factor. For example, Martinet's contrary claim that the likelihood of a sound change depends (other things being equal) on the amount of homonymy it would create has never to my knowledge received more than anecdotal support, and has failed attempts at statistical verification (King 1967). The changes seem to take place anyway, and are then followed by further change which corrects the damage done by the first. That is, language practises "therapy" (Gilliéron) rather than prophylaxis. Discussing the effects of vowel syncope in Odawa, Kaye (1973) comments: "Why a rule should enter the language which simultaneously opacates a stress rule, destroys a surface alternating stress pattern and causes wholesale allomorphy, seems a question worth pondering."

More generally, we could ask: why are changes locally rather than globally motivated? It is likely that the answer lies at least in part in the conventionalizing and spread of originally functionally restricted styles of speech.

The distinct functional targets of change are not necessarily reached in formally distinct ways (as suggested in Vennemann 1972). Any of the three modes of optimalization could in principle be implemented by any possible kind of change. Thus, reordering into feeding order can reduce phonetic complexity, as in nearly all the cases noted in Bailey (1973), e.g. jus' yet becoming juš' yet in faster speech by ordering the deletion of final t in clusters to precede palatalization of s. It can also reduce cognitive (structural) complexity. This will be the case when the fed rule is morphologically conditioned. This happens, for example, in German dialects which apply umlauting to derived back vowels. There are dialects of Swiss German where æi turns into ā before consonants and word-finally (Jutz 1931, 116), e.g. ā 'egg' ~ æiɔr 'eggs', tswā ~ tswæiar 'two-deciliter bottle', šrā 'a scream' šræia 'to scream'. Originally, and still in the dialects of part of the Kanton of Schaffhausen (Wanner 1941), the æi → ā rule precedes umlauting, so that the secondary back vowels it produces are not fronted in the umlaut contexts, as in the diminutive form form āli 'little egg', or in the plural šrā 'screams'. In most dialects, however, the rules are placed in feeding order, with ā from æi undergoing fronting to ǣ in umlauting contexts, giving forms like ǣli in place of original āli (e.g. Enderlin 1913, for the dialect of Kesswil in Oberthurgau). As umlauting is a purely morphological process, its extension to derived back vowels can have no phonetic motivation, but rather reduces the opacity of the umlaut rule. Presumably only the first kind of reordering is favored in fast or careless speech, and only the second can be regarded as an instance of analogical change in the usual sense.

In rule generalization there is again a phonetically and a systemically motivated variety. For example, fast speech brings about a more general application of many rules in English, such as syncope. On the other hand, the generalization of the consonant assimilation rule l+n → l+l in Finnish (e.g. tul+nut → tul+lut) to the form C_i+n → C_i+C_i (e.g. pur+nut → pur+nut, pes+nyt → pes+syt) can hardly be phonetically motivated. The reason is that it does not eliminate Cn clusters from the language, as these are retained morpheme-internally (for further discussion, see Kiparsky 1973a, 71).

The analogous observation can be made also for sound change. Cf. especially Stampe (1972) on polarization and assimilation as, respectively, perceptually and articulatorily motivated phonological processes. One might say that the only limitations on what functions a given formal change can serve are purely logical ones. Thus, rule generalization or reordering into feeding order can scarcely decrease perceptual difficulty, as their effect is generally to suppress contrasts that would otherwise appear on

the surface. And it is hard to think of cases in which it would be possible for the addition of a new rule to make the abstract system more learnable.

Taken together, these considerations imply that there could be no single function of a grammar which would predict the directions of possible change. There are rather several independent factors, whose effect moreover seems to take the form of a loose interplay of local optima.

Although we can generally recognize what sorts of things each of the factors is to account for, a precise explication of any of them is still lacking. Relatively much effort has been devoted to finding out what is involved in learnability (e.g. Brink 1973, Miller 1973, Kaye 1973, Schindler 1974). All recent writings on the subject that I have seen hold that simplicity combined with the current form of grammars does not suffice. More adequate solutions have been sought in two directions: one, taking learnability itself to be determined by several independent subfactors, and the other, seeking to justify reanalyses of grammars and revisions of linguistic theory which have the effect of making simplicity work.

A proposal made in Kiparsky (1971, 1972) was that learnability depends on three things: in addition to (1) formal simplicity, also (2) paradigm coherence, or degree of allomorphy, and (3) transparency of rules. It immediately brings up a problem parallel to the one concerning the mutual relationship of the three functional factors: how are the supposed subfactors that determine learnability in turn related to each other? On the face of it, there are clear cases where they conflict, as would be expected if the factors are indeed truly independent. What are the domains of conflict, and how is it resolved? The answer could be that there is some ranking or partial ranking of factors such that if A ranks over B, then (i) there are changes where A is optimized at the cost of B, and (ii) there are no changes where B is optimized at the cost of A (cf. King 1972 for some discussion). It might also be the case that the resolution is indeterminate, or determined by the grammar or by the language learning situation in some way.

A priori, this whole approach does not seem very satisfactory. Especially the distinction of simplicity and allomorphy is suspect, as Koefoed (1974) has noted: "It sounds like a paradox that a reduction of allomorphy could ever complicate a rule system that has as its main task precisely to account for allomorphy". Some kind of reduction in the components of the evaluation measure would therefore be welcome, especially if it could reduce the questions of the preceding paragraph to pseudo-problems.

Consider, from this point of view, first the principle of minimization of allomorphy. One possible way of getting rid of it would be to claim that given the right theory of grammatical form, minimization of allomorphy is always correctly represented as formal simplification of the system. One could in effect deny

the independence of the notion of allomorphy, if it were the case that simply by virtue of the way grammars work, a reduction in allomorphy (in whatever sense that term appears to be linguistically relevant) must formally simplify the grammar, and any increase in allomorphy must correspondingly complicate it. (The converse would not necessarily have to be true). Koefoed notes that a large step can be taken towards this situation by redefining "degree of allomorphy" in terms of alternation types rather than alternation tokens. Then, the degree of allomorphy is given by the number of alternating pairs of segment types rather than by the number of pairs of allomorphs exhibiting such alternations. A consequence of this is to make allomorphy dependent in most cases just on the presence of a rule, not on its generality. Given this definition of degree of allomorphy, Koefoed claims, simplification in the rule system never leads to new alternation types. He excepts only cases where this simplification is motivated "by phonetic factors rather than by the abstract tendency to formally simpler rules". That is, new alternation types can have phonetic but not systematic causes. But even with this proviso, the claim is not strictly true. The Swiss German and Finnish examples mentioned earlier are typical counterexamples. Koefoed could of course respond to this objection by defining "alternation type" at a still higher level of abstraction, in terms of classes of segments, or features. On this definition, the cited reordering of the umlaut rule does not produce a new alternation type, but merely extends an old one, viz. the alternation between front and back in the vowels. Some further redefinition would be required for rules involving deletion, total assimilation (as in the Finnish example), and so on.

But this would still be insufficient, for it leaves allomorphy frequently failing to correspond to formal complexity. This happens, on the one hand, where a rule accounts for more than one surface "alternation", and on the other, where more than one rule accounts for a single paradigmatic target.

A single rule can account for more than one alternation when its output is differentiated by rules that apply after it. Because of this possibility, rule generalization can increase allomorphy in the sense of creating new alternations, if these are defined as types of differences between related surface forms. An example is Kuryłowicz's (1968, 298) derivation of IE lengthened grade by the proportion RT:eRT = eT:ēT. This corresponds to a generalization of an inherited ablaut rule ($\emptyset \rightarrow e$) from $\overset{\circ}{C}R\overset{\circ}{C}$ stems ($\overset{\circ}{R}T \rightarrow eRT \rightarrow eR\overset{\circ}{T}$) to CVC stems ($eT \rightarrow eeT \rightarrow \bar{e}T$), creating a new type of alternation in length by analogy to an existing one that involves the presence or absence of a short vowel. Since allomorphy is by definition a relation involving surface representations, such cases show that allomorphy cannot possibly translate point by point into complexity, as required by the proposal we are considering.

The opposite type of case, where more than one rule converges on a single allomorphic target, likewise falsifies the

strict reducibility thesis. A uniform paradigm could, given the present theory of grammar, easily be due to several rules which separately would not produce this result. In such cases, the uniform paradigm could formally arise through the addition of a rule (a complication of the system), and correspondingly, an elimination of one of the rules (a simplification) could introduce new allomorphic variation into surface structure. Not only is this a logical possibility, but analyses have been proposed in which just the described situation holds. Apparent systemic complications resulting in paradigm leveling under such conditions were indeed the original motivation for the proposal that a principle of minimization of allomorphy be introduced as part of the evaluation measure in addition to simplicity. The standard examples include phenomena like columnar stress in inflectional paradigms, which in generative analyses of several languages, e.g. Spanish (Harris 1973) and Sanskrit (Kiparsky 1973b) mysteriously arises from an elaborate set of conditions of a morphological sort on the stress rule, or even from the interaction of several stress rules, and the leveling of the original s/r alternation in the Latin nouns of the type honor, where the change from honōs:honōris to honor:honōris seemed to require a new branch of rhotacism in the Nominative Singular. Since these are at least possible analyses, the claim that allomorphy is necessarily equivalent to formal complexity evidently breaks down here.

The remaining possibility of eliminating a separate principle of minimization of allomorphy is more promising. This is to find a theory of grammar in which it will be true that actually occurring types of leveling will always constitute formal simplifications. It involves conceding the formal point that leveling could involve a complication of the grammar, and replacing it with the empirical claim that it in fact never does. It also involves withdrawing the claim that simplification could never result in an increase of allomorphy, which, as we have seen, is dubious anyway. Leveling would then simply be a reflex of some types of simplification, rather than a factor in its own right.

What is interesting about this version of the proposal to eliminate a principle of elimination of allomorphy from the evaluation measure is that it restricts leveling to cases where it is associated with a formal simplification of the system. Other, theoretically imaginable types of leveling accordingly should not be attested in actual change.

The crucial test cases are precisely those in which leveling has been analyzed as resulting from a formal complication in the grammar. If the proposal is to go through, reanalyses must be motivated for these examples, in which the relative formal simplicity of the stages before and after the changes is reversed. It is significant, then, that a more detailed analysis by Harris (1974) of the development of columnar stress in Spanish has indicated that the steps in this development constitute progressive simplifications of the stress rule that arose when the quantity distinction of Latin was first lost. In the Latin example, it seems reasonable to suppose that the change of honōs to honor involves

not a new environment for the rhotacism rule, but simply a reanalysis of -ōs stems to -ōr stems. This hypothesis would at once explain the restriction of the leveling to masculine and feminine polysyllables, for it is precisely in these categories that stems in -ōr existed at all (Schindler 1974, who further refers to Szemerényi). Recently Redenbarger (1974) has argued that in the context of the Latin morphophonemic system the change reduces to a reordering of two rules.

The evidence for an independent principle of minimization of allomorphy that has so far been presented is thus rather poor. However, the decision to abandon it should not be made lightly, as it has some surprising consequences. Vennemann (1973) has pointed out, in connection with motivating his theory of Natural Generative Grammar, that the reduction of leveling to formal simplification is often possible only if phonological representations are assumed to be more superficial than those countenanced in standard generative phonology. A particularly clear instance is the change in Canadian English from the [rəyt]:[raydr̥] type of dialect to the [rəyt]:[rəydr̥] type, thoroughly discussed by J.K.Chambers (CJL 1973). The effect of this change is to eliminate the a:ə alternation arising in diphthongs through the well-known interaction of voicing with the "Canadian Raising" rule. As far as I can see, this must be considered a case of analogical leveling, and there is no other reasonable motivation for the change. Yet it can be represented as a formal simplification only if we assume that the raised diphthongs prior to the change are lexicalized before underlying voiceless stops, as in write /rəyt/, with a lowering rule then applying before voiced consonants to give forms like writer [raydr̥]. The change from the type [raydr̥] to the type [rəydr̥] would then be the loss of the lowering rule. Current generative phonology would however require the redundant difference between [ay] and [əy] to be taken out of underlying representations and introduced by a phonological rule instead.

Turning now to the question whether opacity might be similarly eliminable as a separate factor in the evaluation measure, we face the same two possible strategies as in the case of allomorphy, except that the difficulties this time turn out to be overwhelming. One possibility, again, is that under the right formulation of grammar opacity might turn out to be dispensable because it always translates into complexity. Any increase in opacity would then have to formally complicate the grammar, and any decrease in opacity would have to simplify it. But this is hardly ever so. For one thing, every case of rule reordering is a counterexample, as it causes a decrease in opacity with no corresponding change in formal complexity of the grammar. True, disallowing extrinsic ordering (Koutsoudas, Sanders, and Noll 1974) generally has the effect of forcing a reformulation of non-feeding ordering by some device which complicates the system (e.g. an extra restriction on the fed rule). What the standard theory would represent as reordering into feeding order (necessarily a gain in input transparency) would then constitute a formal simplification. This would eliminate the

need for an independent factor of opacity in certain cases. Aside from the problems that the exclusion of extrinsic ordering meets with (e.g. Campbell 1973) it remains true that outside of this class of disputed cases, opacity does not (even under the no-ordering theory) correspond to complexity as far as we can now tell. This is true for reordering into bleeding order, which brings a gain in environment transparency. The Koutsoudas-Sanders-Noll theory of rule ordering does not guarantee the correct directionality in such cases. Until a principled basis is given for translating environment opacity into complexity, the elimination of extrinsic ordering does not do away with the need for opacity.

Indeed, to find cases where opacity does not correspond to formal complexity we need not turn to rule ordering at all. The Finnish case cited above, where a rule is generalized and becomes opaque in the process, illustrates the basic independence of the two concepts.

Still less acceptable is the second alternative possibility of eliminating opacity, whose counterpart turned out to be rather promising in the case of allomorphy. This would be to give up the equivalence of opacity with complexity, which fails in such obvious ways, and to rather claim that opacity can be reduced to simplicity in those cases where it is apparently relevant to linguistic change. A review of the cases we have just mentioned quickly refutes this try too. In precisely those cases where opacity is completely independent of complexity its historical effect is most clearly apparent, as for example in reordering of rules. Moreover, opacity seems to play a role in the theory of loan phonology and in the theory of exceptions (Kiparsky 1971). The postulation of a separate principle of opacity in the evaluation measure seems unavoidable.

As to the relative weight of simplicity and opacity where the two principles conflict, the only thing that can be said with certainty so far is that simplification can take place in such a way as to create opacity. We have mentioned an example from Finnish which shows this possibility. The question whether a hierarchical relationship holds, with simplicity dominating transparency, therefore depends on whether opacity is ever eliminated at the price of complicating the grammar.

Our conclusions can be summed up as follows:

(1) Minimization of allomorphy, of opacity, and of complexity are formally independent, in the sense that none of them can be defined in terms of any of the others.

(2) No unequivocal evidence exists for considering minimization of allomorphy as a separate motivating factor in linguistic change. Actually attested cases of leveling can usually with some plausibility be represented as formal simplifications, though this in some cases involves nontrivial and perhaps false assumptions about the form of grammars.

(3) Minimization of opacity must, however, be considered to function as a separate motivating factor in linguistic change, and hence (by our assumption about the source of such change) also as a separate component of the evaluation measure.

The idea ths
at least two separate
have reached on stric
generalized naively s
a corpus of examples
what the system is li
of difficulty for the
ion which must be pro
random digits is hard
of memory space. A li
morphology. There is
which is for example
to "learn" codes. A
a hard-to-figure-out
it might be easy to
covered. As Koefoed (
rule might be conside
so far proposed also
output. Even a single
discover and might th
example the rule for
this fact which is th
difficulty would of
applying sequentially
of learning difficult
as complexity and ope
measure of the amount
stored. Opacity would
the ease with which
ered (which is not to
an effect on it). Per
point where linguist
can finally begin to
psychology.

Bailey, Charles-James
ferent rule
James Baile
Analyzing V
D.C.:George
Brink, Daniel (1973)
application
Amsterdam.
Campbell, Lyle (1973
report No.
University
Enderlin, F. (1913)
Beiträge zu

The idea that the "learnability" of a system depends on at least two separate factors, complexity and opacity, which we have reached on strictly language-internal evidence, might be generalized naively as follows. Learning an abstract system from a corpus of examples involves two distinct tasks: discovering what the system is like, and fixing it in one's memory. A source of difficulty for the latter task is the sheer amount of information which must be processed and stored. Thus, a long list of random digits is hard to learn simply because it takes up a lot of memory space. A linguistic counterpart would be suppletive morphology. There is another source of difficulty, of the sort which is for example maximized in cryptography, to make it hard to "learn" codes. A formally very simple set of rules can produce a hard-to-figure-out relation between input and output, though it might be easy to remember the rules once they have been discovered. As Koefoed (1974, fn. 4) notes, the transparency of a rule might be considered to depend in addition to the factors so far proposed also on the phonetic similarity of its input and output. Even a single, formally simple rule can be hard to discover and might therefore be considered transparent, as for example the rule for turning English into Pig Latin (it is just this fact which is the whole point of using Pig Latin). The difficulty would of course be compounded by several such rules applying sequentially. It is natural to identify these dimensions of learning difficulty with what we have identified linguistically as complexity and opacity. Complexity could be thought of as a measure of the amount of information which must be processed and stored. Opacity would then be a measure of one factor determining the ease with which the structure of the system may be discovered (which is not to deny that complexity would not also have an effect on it). Perhaps this speculation can take us to a point where linguist's hypotheses about language acquisition can finally begin to make contact with findings of experimental psychology.

References

Bailey, Charles-James N. (1973). "Variation resulting from different rule orderings in English phonology". Charles-James Bailey and Roger W. Shuy (edd.), New Ways of Analyzing Variation in English, pp. 211-52. Washington, D.C.:Georgetown University Press.

Brink, Daniel (1973). "Characterizing the natural order of application of phonological rules" (MS), Berkeley/Amsterdam.

Campbell, Lyle (1973). "Extrinsic ordering lives". Technical report No. 80, Center for Research in Social Behavior, University of Missouri, Columbia.

Enderlin, F. (1913). Die Mundart von Kesswil. A. Bachmann (ed.) Beiträge zur schweizerdeutschen Grammatik, Vol. 5.

Ferguson, Charles A. (1971). "Absence of copula and the notion of simplicity:a study of normal speech, baby talk, foreigner talk, and pidgins". Dell Hymes (ed.) Pidginization and creolization of languages, p. 141-50. Cambridge: University Press.

Jutz, L. (1931). Die alemannischen Mundarten. Halle (Saale).

Harris, James (1973). "On the order of certain phonological rules in Spanish". Stephen R. Anderson and Paul Kiparsky (edd.) A Festschrift for Morris Halle, pp. 59-76. New York: Holt.

Harris, James (1974). "Stress assignment rules in Spanish". Read at the Texas Symposium on Romance Linguistics , March 1974. (Proceedings to be published by the University of Texas Press).

Kaye, Jonathan Derek (1973). "Opacity and recoverability in phonology" (Ms.).

King, Robert D. (1967)."Functional load and sound change", Lg. 43. 831-52.

King, Robert D. (1972). "A note on opacity and paradigm regularity", LI 3.535-8.

Kiparsky, Paul (1971)."Historical linguistics". W. Dingwall (ed.) A survey of linguistic science, p. 576-649. University of Maryland.

Kiparsky, Paul (1972). "Explanation in phonology". S. Peters (ed) Goals of linguistic theory, pp. 189-227. Englewood Cliffs: Prentice-Hall.

Kiparsky, Paul (1973a). "Phonological representations". O.Fujimura (ed.) Three dimensions of linguistic theory. Tokyo: TEC Co.

Kiparsky, Paul (1973b). "The inflectional accent in Indo-European", Lg. 49.794-849.

Koefoed, Geert (1974). "On formal and functionalexplanation: some notes on Kiparsky's 'Explanation in phonology'". J.M. Andersson and Charles Jones (edd.) Proceedings of the First International Conference on Historical Linguistics. Dordrecht:Reidel (in press).

Koutsoudas, Andreas, Gerald Sanders, and Craig Noll (1974). "The application of phonological rules", Lg. 50.1-28.

Kuryłowicz, J. (1968). Indogermanische Grammatik, Vol. II. Heidelberg:Winter.

Miller, D. Gary (1973). "On the motivation of phonological change". B.B. Kachru et al. (edd.) Papers in honor of Henry and Renee Kahane. Urbana: University of Illinois Press.

Redenbarger, Wayne J. (1974). "Vowel lowering and i-epenthesis in classical Latin". Read at the Texas Symposium on Romance Linguistics, March 1974 (proceedings to be published by the University of Texas Press).

Schindler, Jochem (1974). "Fragen zum paradigmatischen Ausgleich". Die Sprache, vol. 20 (in press).

Stampe, David (1972). "On the natural history of diphthongs". CLS Papers, Vol. 8, pp. 578-590.

Vennemann, Theo genannt Nierfeld (1973). "Phonological concreteness in Natural Generative Grammar" (MS).

Vennemann, Theo (1972). "Phonetic analogy and conceptual analogy". Theo Vennemann and Terence Wilbur (edd.) Schuchardt, the Neogrammarians, and the transformational theory of phonological change. Frankfurt:Athenäum.

Wanner, G. (1941). Di
(ed.), Beiträge

Wanner, G. (1941). Die Mundart des Kantons Schaffhausen. A. Bachmann (ed.), Beitr#ge zur schweizerdeutschen Grammatik, Vol. 20.