Session 8:   INFORMATION PROCESSING AND LINGUISTIC ANALYSIS

FROM MODEL TO PROCEDURE

Paul L.   Garvin[1]

Ramo-Wooldridge and Wayne State University

A model can be a frame for the acquisition of knowledge,   or a frame for the presentation of results.   I am interested in the former.

Several models have recently been proposed in linguistics which involve an "as if" mode of thinking.   An "as if" model in effect suggests: "Since we cannot deal with our object of cognition effectively or directly,   let us look at it as if it were something else with which we can deal,   and which is sufficiently similar to it so that we can generalize back".   Thus, let us look at the grammar of a language as if it were a probability matrix or a sentence-generating machine.

A variant of the "as if" model is the reductionist model,   in which certain properties of the object of cognition are eliminated because they are considered intractable.   It is thought that this reduces the complexity of the object.   In linguistics, a common application of this reductionism is the attempt to eliminate considerations of meaning from the study of language,   which is then defined as being form only,   or as being describable in terms of form only.

What is studied,   then,   is not the object itself but the particular model.   Such an approach is necessary if the object of cognition indeed does not lend itself to direct investigation.   I contend that this is not so in the case of language.

It is perfectly possible to look upon language the way we have observed it to be in our past experience and to systematize this observation by abstracting from it a set of assumptions about the characteristic properties of our object of cognition,   such that it is possible to study the object in terms of these properties.   If we formulate these properties,   we obtain,   not a model in the previous sense,   but a definition in the classical sense of the word.   Such a definition will, however,   serve as a perfectly good frame for the acquisition of knowledge; and why not,   therefore,   consider it a model in its own right and call it a "definitional model"?

---

[1] Editor's note: Dr.   Garvin was a member of the faculty of Georgetown University   at the time of the Symposium.

Session 8: INFORMATION PROCESSING AND LINGUISTIC ANALYSIS

To be useful, the properties set forth in a definitional model must be general enough not to prejudge the results, and specific enough not to be trivial.   They must allow the differentiation of one's object of cognition from other similar or related objects, and they must provide suitable points of departure for a procedure.   They must allow the deduction of the equivalence principle and relevance criterion proper to one's own field of study.

I stated the essentials of such a definition of language in 1948 in a paper before the 29th International Congress of Americanists;[2] the analytic work which has led me to this definition and has followed from it has not invalidated it.

As all classical definitions, my definition contains a genus and a differentia.   It places language into the larger class of phenomena to which it belongs by stating that it is a system of signs.   It differentiates it from other phenomena in the same class by stating that language differs from other systems of signs by a particular set of structural properties.

These properties can be formulated in terms of three sets of levels: two levels of structuring, the phonemic and morphemic, respectively; two levels of organization, namely selection and arrange-ment; and several levels of integration, along which the scale of units of increasing complexity is arranged.

Note that my definition of language as a system of signs does not include the term "vocal".   The reason is that the vocal character of spoken language is an irrelevant substantive property, since it is a mode of manifestation rather than an essential attribute of the system. To limit language to this one mode of manifestation would violate the requirement of generality for the definition.

Note also that in stating levels of integration I have not specified a particular number of levels, but merely that there be several levels, in order not to impose an unnecessary specificity of structure on any particular language.

The three sets of levels are intended to differentiate the language which is the linguist's object of cognition from other systems, often

---

[2] "Structure and Variation in Language and Culture", Proceedings of the 29th International Congress of Americanists, New York, 1948 (Chicago, 1952), vol. 3.

called languages,   which are not.   Thus,   when logicians talk about a
simplified language as defined,   for instance,   by the vocabulary "a",
"b",   "c", . . .   and the syntax "+",   "-",   "=",   this is not a language in
our sense,   since it lacks,   of the three sets of levels,   both the required
two levels of structuring and the levels of integration.   Similarly,
the language of the bees as described by von Frisch is not a linguist's
language,   since it,   too,   is limited to levels of organization only.

From the properties stated in the definition can be deduced a
set of methodological principles on which procedures can be based.

The generic part of the definition,   namely,   that language is a
system of signs,   allows me to posit the association of form and
meaning through the sign function.   From this association follow the
equivalence principle and relevance criterion proper to linguistics.

The linguistic equivalence principle will allow differentiation
between what is same and what is not same in linguistics: same is
what is functionally equivalent,   and not necessarily what is sub-
stantively identical.   Allomorphs,   for example,   have different forms
but the same meaning--they are substantively not identical,   but
functionally equivalent.   Conversely,   homonyms are substantively
identical,   but functionally not equivalent.   Thus,   only if form and
meaning are considered together can sameness and difference be
established.   A consequence of the linguistic equivalence principle is
the linguistic relevance criterion: that which affects functional
equivalence is relevant; that which does not affect functional equival-
ence,   although it may affect substantive identity,   is not relevant.

The phonemic and morphemic levels of structuring differ in the
type of the association of form and meaning: phonemes are meaning
differentiators,   morphemes are meaning carriers.   Consequently,   in
morphemics the association of form and meaning is in the nature of a
covariance,   in phonemics it is not.   This is operationally of the great-
est significance: where there is covariance,   one of the covariants can
be made into the independent variable,   whereby the other covariant
becomes the dependent variable.   This is what happens in the elicita-
tion of paradigms,   which is one of the standard techniques for han-
dling morphemic data.   In phonemics,   on the other hand,   since there
is no covariance,   paradigms cannot be elicited but can only be
compiled ex post facto.

The two levels of organization, which are stated in terms of the relevance of arrangement in addition to selection,   allow not only the manipulation of paradigms as obtained above to establish the membership of each class,   but also the manipulation of sequential order which leads to the differentiation of classes from one another.

The levels of integration imply the existence of not only minimum units,   namely phonemes and morphemes,   but also of fused units,   as I have discussed in some detail in my paper in Session 6 of this Symposium.   For each fused unit there can be posited an internal structure and an external functioning as separate attributes.   The separateness of these attributes allows the application of the two basic procedures of dropping and substitution to fused units in morphemics.   In dropping, portions of sequences are omitted from the whole,   leading to the establishment of relations of dependence,   which are one mode of external functioning.   Classes of fused units can be defined in terms of these relations.   In substitution,   units of different internal structure are substituted for each other in controlled frames,   such as the ones resulting from the above relations,   within which they have the same external functioning.   Inventories of classes can be compiled in terms of this substitutability.

In actual analytic practice, procedures stemming from different properties of the model are used in order of their applicability to the problems encountered,   rather than in the order in which the model was presented.   They do, however,   fall into two broad classes: procedures for the establishment of primary units,   that is,   initial cutting, and procedures for the establishment of distributions.   The former draw, in some form or other, upon the association of form and meaning.   The latter do not.   Those procedures which stem from this association are the linguistic analog of psychological experimentation, based on a rigorous control of the responses of the subject,   that is, the native speaker.   The non-psychological procedures consist in a logical manipulation of the results of the former.

From the standpoint of data processing,   only the logical procedures in linguistic analysis are computable; the psychological procedures are not,   but their results are essential as the input to the program which can be written to implement the former.