ELIMINATING FORMAL CONSTRAINTS OF LANGUAGE

Yu. S. Martem'janov

**All-Union Centre for Translation
of Scientific and Technical
Literature and Documentation**

Abstract

In order to eliminate the formal constraints of language from a natural-language text, a series of artificial languages are needed to detect the various types of constraint.

   In order to eliminate the formal constraints or conventions
of a text in a natural language (NL), an artificial language
(AL) is needed in which any expression, while being a faithful
representation of a natural text with regard to both <u>content</u>
and <u>conventions of expression</u>, would make a clear and explicit
distinction between these two aspects.

   If an AL is to be used for distinguishing linguistic
conventions, its grammar should have the following features:

   a) a 'modulation' component for transforming the basic
      expressions in the canon in order to add certain
      features of various types, and
   b) a distinction between 'intra-linguistic' features and
      others, for example concord between the morphological
      endings of interdependent words ('agreement' and
      'rection'), the predicative status accorded to certain
      constituents of the text (propositional arrangement) or
      the articulation of a linguistic complex of unusual
      depth (in the sense used by Yngve) in an arborescent
      pattern of fragmentary propositions, etc.

   For any AL expression using a canon modulated in accord-
ance with the intralinguistic features (and which thereby
coincides with a given natural text), such a grammatical
structure makes it possible to recognise these features and
eliminate them without compensation. The resulting "de-idio-
matized' AL expression thus represents the deep structure of
the natural text, approximating to its extra-linguistic content.
The translation to the deep structure is accompanied by the
elimination of lexical conventions: the variable words which
serve only to express the intra-linguistic features (agreeing
and subordinate forms, tool-words for propositional arrangement
or anaphorization, etc.) are replaced by corresponding
invariable lexemes.

   The intra-linguistic features are not homogeneous with

regard to the presumed grammatical categories: thus, morpho-
logical agreement presupposes a syntax of 'parts of speech'
(such as 'noun', 'verb', 'preposition'); a change in propo-
sitional status (e.g. changing the subject into an adverbial
phrase: The use of semi-conductors simplifies assembly - With
the use of semi-conductors assembly is simplified) necessitates
a non-propositional syntax of valences and complements; finally,
the expression of an initial linguistic complex in given forms
presupposes that the syntax is dictated solely by semantic
relationships ('cause', 'agent', etc.) linking the necessarily
elementary arguments. Thus three different ALs are needed,
each with its own base and appropriate modulations - i.e., in
order of abstraction: the 'verbal' AL containing the 'parts of
speech', the 'valence' AL (see above) and the 'relationships'
AL, with a syntax of semantic relationships

The proposed ALs form a series, starting with the verbal
language, which coincides in its surface structures with the
propositions in NL, and ascending towards more abstract ALs.
Once the linguistic peculiarities have been picked out using a
particular AL, the deep structure thus obtained can be trans-
lated into a surface formula in a more abstract AL which permits
further de-idiomatization until we arrive at the deep structure
of the 'relationships' AL, which is free of all linguistic
conventions.

The NL text, inserted at the verbal end of the proposed
series of AL can thus be transformed into a deep structure in
any AL, depending on the degree of de-idiomatization required.