

Sentence-for-sentence translation*

Victor H. Yngve, Research Laboratory of Electronics and Department of Modern Languages, Massachusetts Institute of Technology

Introduction

Recent advances in linguistics, in information theory, and in digital data-handling techniques promise to make possible the translation of languages by machine. This paper proposes a system for translating languages by machine — with the hope that when such a system is worked out in detail, some of the language barriers can be overcome. It is hoped, too, that the translations will have an accuracy and readability that will make them welcome to readers of scientific and technical literature.

Word-for-word translation could be handled easily by modern data-handling techniques. For this reason, much of the work that has been done up to this time in the field of mechanical translation has been concerned with the possibilities of word-for-word translation^{2,3}. A word-for-word translation consists of merely substituting for each word of one language a word or words from the other language. The word order is preserved. Of course, the machine would deal only with the written form of the languages, the input being from a keyboard and the output from a printer. Word-for-word translations have been shown to be surprisingly good and they may be quite worth while. But they are far from perfect.

Some of the most serious difficulties confronting us, if we want to translate, arise from the fact that there is not a one-to-one correspondence between the vocabularies of different languages. In a word-for-word translation it is necessary to list alternative translations for most of the words, and the choice among them is left up to the ultimate reader, who must make his way through a multiple-choice guessing game. The inclusion of multiple choices confuses the reader or editor to the extent that he is unduly slowed down, even though he can frequently glean the correct meaning after study. Another great problem is that the word order — frequently quite

different in the two languages — further obscures the meaning for the reader. Lastly, there are the more subtle difficulties of idioms and the particular quaint and different ways that various languages have of expressing the same simple things. While it has been suggested in the past that rough word-for-word translations could be put into final shape by a human editor, the ideal situation is that the machine should do the whole job. The system proposed here is believed to be capable of producing translations that are considerably better than word-for-word translations.

The solution of the problems of multiple meaning, word order, idiom, and the general obscurity of the meaning when translation is carried out on a word-for-word basis is to be found in translating on a sentence-for-sentence basis. Nearly all of these problems can be solved by a human translator on a sentence-for-sentence basis. By this we mean that each sentence is translated without reference to the other sentences of the article. This procedure can be simulated experimentally by separating a text into sentences and submitting each for translation to a separate person who would not have the benefit of seeing any of the other sentences. In most instances an adequate translation of each sentence would result. Very little would be lost by discarding all of the context outside of one sentence length.

There are striking parallels between language and error-correcting codes. Language is a redundant code, and we are here proposing to deal with code blocks longer than one word, namely, with blocks of a sentence length. Our problem is to specify the constraints that operate in the languages out to a sentence length. This will be difficult because languages are so complex in their structure. However, we shall attempt to specify these constraints, or at least to lay the foundation for such a specification.

The Nature of the Process

A communication system may be looked upon as having a message source, an encoder, a statement of the rules of the code or a codebook for encoding, a decoder, a statement of the rules of the code or a codebook for decoding, and a destination. (See Fig. 1.) The function of the message source is to select the message from among the ensemble of possible messages. The function of the rules of the code or the codebook

*This paper was presented at the Third London Symposium on Information Theory, September 12 to 17, 1955. A shortened version with discussion will be published in the proceedings of the conference under the title *Information Theory* by Butterworths Scientific Publications in 1956. An earlier version of some of the ideas contained in this paper can be found in Chapter 14 of reference 2. This work was supported in part by the Signal Corps, the Office of Scientific Research (Air Research and Development Command), and the Office of Naval Research; and in part by the National Science Foundation.

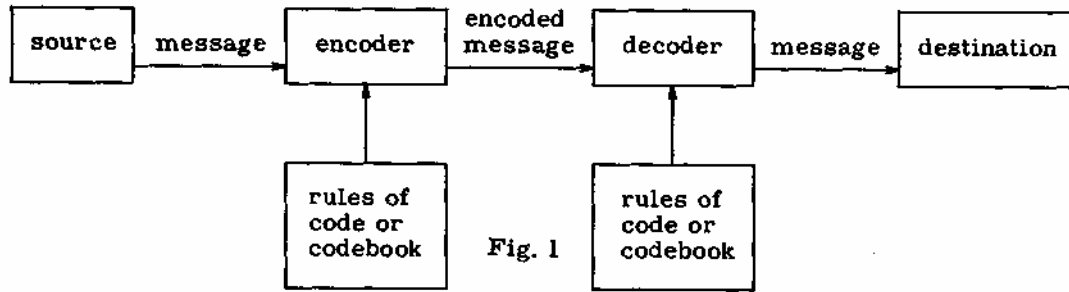


Fig. 1

is to supply the constraints of the code to which the encoded message must conform. In general, the encoded message is in a more redundant form than the original message. The function of the decoder is to recognize the features of the encoded message that represent constraints of the code, remove them, and supply the destination with a message that is a recognizable representation of the original message. This characterization of a communication system can be used with advantage to represent language communication only if great care is used in interpreting the various concepts. To this we shall now turn our attention.

In the case of language communication there is no difficulty in specifying what is meant by the concept of an encoded message if we restrict ourselves to the conventional written representations of the languages. Such written representations can be expressed in binary or other convenient form. What we might mean by "message," however, is very difficult to specify exactly. Here we encounter some of the many difficulties with "meaning" that have plagued linguists. In the first place, it is very difficult to separate a message source from an encoder when the same individual performs both tasks. The message here would be, approximately, some representation of the "meaning" that the individual could express in the different languages that he might know; it would be something common to all of the different language representations. The message that arrives at the destination would be the receiver's under-

standing of the meaning, and might not, in fact, be the same as the message that left the source, but usually it is approximately the same if the individuals using the language understand each other. The decoder might not recover the original message, but another, and then there would be a misunderstanding. The decoder might extract a message quite different from the one intended by the message source, as a result of a confusion between message and constraints, and this might happen if the rules used by the decoder are not exactly equivalent to the rules used by the encoder. In this case, some of the constraints supplied by the encoder might not be recognized as constraints by the decoder, but interpreted instead as part of the message. For example, the encoded form of the message might be "Can you tell me where the railroad station is?" and the decoder might extract such a message as "This person speaks English with an American accent." Or, as another example, the child who receives encoded messages in a language gradually accumulates information about the rules of the language and how to use it.

We now shift our attention from communication systems employing a single code or language, to systems which translate from one code or language into another. A code translation system can be looked upon as being much the same as the above representation of a communication system, but with the operations carried out in a different order; the positions of the encoder and the decoder are reversed. (See Fig. 2.) If the

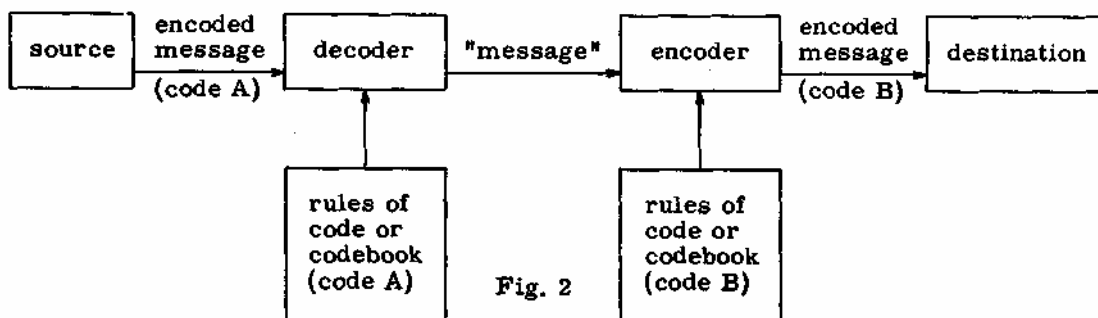


Fig. 2

codes are very similar, or in some sense equivalent, it may not be necessary to first decode and then encode. It may be necessary only to partially decode. If the two codes are very different, it may be simpler to decode to a minimally redundant form of the original message before encoding in the new code. We would like to consider the process of language translation as a two-step process: first, a decoding, or at least a partial decoding; then a recoding into another of the hundreds of known languages. The difficulties associated with word-for-word translations arise from the use of only a partial decoding, that is, a decoding based on the word instead of the sentence or some larger block.

We can assume that most material in science and engineering is translatable, or expressible in all languages of interest. An expression and its translation differ from one another in that they conform to the different constraints imposed by two languages. They are the same in that they have the same meaning. This meaning can be represented by some less redundant expression that is implicit in both language representations and that can be obtained by stripping off from one of them the trappings associated with that particular language. This representation might be called a transition language. Attempts at a specification of the structure of the "message" may get us into some of the difficulties associated with "meaning" but a description of the same thing as a transition language comes naturally from a description of the constraints of the two languages, since the transition language is just a representation of the freedom of choice left after the constraints of the languages have been taken into account.

Many of the constraints of language are quite constant. Grammar and syntax are rather stable. But there are other constraints that are peculiar to each user of the language, each field of discourse, each cultural background. A restriction can perhaps be made in mechanical translation to one field of discourse so that it will be easier to specify the constraints. Since language is a very complicated coding system, and in fact not a closed system, but an open one in that new words, constructions, and innovations are constantly being introduced by various users, the complete determination of the constraints is practically impossible. The best that one can do is to determine an approximate description of the constraints that operate; thus our translations will remain approximate.

What we mean by the concept of transition language in a language translation process can be illustrated by the word-for-word translation case. Booth⁴ pointed out that one could not go directly from the words of one language to the words of another language with a digital computer of reasonable size, but that it would be more economical to go through the intermediate step of finding the addresses of the output words. These addresses are in a less redundant form than the original words, and for the purpose of this discussion they will be considered as the transition language. What we mean by transition language in a mechanical translation process is the explicit directions for encoding which are derived by the decoder from the incoming text.

The practical feasibility of mechanical translation hinges upon the memory requirements for specifying the rules of the code, or the structure of the languages. Word-for-word translation is feasible because present-day digital data handling techniques can provide memories large enough to store a dictionary. In other words, we can use a codebook technique for decoding and encoding on a word-for-word basis. If we want to translate on a sentence-for-sentence basis, we must find some method for specifying the structures of the languages which is compact enough to fit into practical memories. Obviously we cannot extend the dictionary concept by listing all of the sentences in the language with their translations. There are certainly in excess of 10^{50} sentences less than 20 words in length in a language like English.

Our problem, then, is to discover the constraints of the language so that we can design practical encoders and decoders. Our problem is that of the linguist who would discover such constraints by careful observation of encoded messages. The following example from coding will illustrate some important aspects of the problem of discovering constraints. We are given the data that the following four binary digit sequences are some of those allowed in the code. We are to determine the constraints of the code.

```
10101010    01001011
11100001    01100110
```

Here, as in the case of studying the structure of language, we do not have an exhaustive list of the allowed sequences. We can only make tentative hypotheses as to the exact form of the constraints and then see if they predict the existence of other observable sequences. Thus we might guess that one of the constraints in the

code above is that the number of 0's and 1's is the same. The hypothesis will fall as soon as the sequence 0000000 is observed. Of course the linguist would make short work of the simple coding problem and would soon discover that there are only 16 different allowed sequences. If he were clever, he might deduce the rules of the code (the structure of the language) before he had obtained samples of all of the sequences. He might discover that the second four digits are identical with the first four digits if there is an even number of 1's in the first four; and that if the number of 1's in the first four digits is odd, the second four digits are the complement of the first four, formed by replacing 0's with 1's, and 1's with 0's. Having this specification of the rules of the code, he can say that it takes four digits to specify the message, the other four being completely determined by them. He might then say that we can take the first four digits as the message. He could equally well have chosen any four independent digits, such as the last four, or the middle four. This corresponds merely to assigning to the 16 messages 16 numbers in different order. The code has error-correcting properties, as does language. If one of the eight digits is in error, its location can be deduced by comparing the first four digits with the last four digits, and checking the parity of the first four. If there are two errors, either the first and last four digits differ in two places, or there are no differences, and the parity of the first four digits is odd.

The solution to our little coding problem is satisfactory in that we have a very compact statement of the constraints of the code. However, if we want to utilize the code in an actual communication channel, we have to design an encoder and a decoder. It may be that there are other simple statements of the rules that might be more suitable for the processes of encoding or decoding. In fact, there are other such representations, since the code above is equivalent to the Hamming code⁵ of this length, for which the rules for encoding and decoding can be stated entirely in terms of parity checks. The code is also equivalent to the Muller-Reed code^{6,7} of this length which uses a majority rule test in decoding. The three statements of the rules of the code are all valid. The choice of the representation of the rules of a language depends partly upon the use for which it is intended, and it is quite possible that one choice would be made for use in encoding and another choice would be made for use in decoding. In

other words, the rules of a language may be phrased in a number of equivalent ways. For use in translating machines, they must be operational, that is, they must be appropriate for use in a machine that operates by a pre-determined program⁸.

The coding example given above illustrates five points about the language problems connected with mechanical translation. First, the rules of the code must be determined from an examination of the received messages. Second, there is no unique specification of the message. Third, there is redundancy which is useful for error correction. Fourth, there may be many equivalent formulations of the rules of the code. Fifth, the choice of a formulation depends partly upon the use for which it is intended.

If our purpose is translation, there is one further consideration. The choice of the form of the rules is also dependent upon which two languages are involved in translation and also in which direction translation is being carried out. It is very likely that the rules of English will have to be restated in various forms, depending on whether one wants to translate into German, out of German, into Russian, out of Russian, and so on. The reason is that certain relations can be found between different languages which can be used to simplify the process of decoding and encoding for the purposes of translation. The form of the transition language that forms the intermediate step in translation will be different with different language pairs.

We have pointed out that we want to translate on a sentence-for-sentence basis; that the feasibility of being able to do this depends upon whether or not we can state the structures of the languages in a form that is sufficiently compact for storing in a machine memory; and that the form of the statements of the structures must conform to certain other requirements, chief among them being that they be appropriate for use in decoders and encoders. We now proceed to discuss the problem of specifying language structure for use in mechanical translation processes.

Structure of Language from the Point of View of the Encoder

We want to consider, first, the form of the rules from the point of view of the encoder because they are simpler to explain and correspond more

closely to other points of view commonly encountered. The encoder combines the message with the rules of the language in order to form the encoded message.

We want to limit the encoder to the words of the language. Of the various ways of doing this, perhaps the only one that seems feasible is to list the words of the language in a dictionary and to store this dictionary in the machine. Whether or not an attempt is made to reduce the number of entries in the dictionary by the use of a stem-affix routine — as is proposed by several authors — or by a method of splitting up compound words⁹, depends upon whether it will be more economical to supply the required routine or to supply the additional storage space needed to list in full all of the words in their various inflected forms.

We want to encode in blocks of a sentence length. Since the words are to be listed in a dictionary, it seems appropriate to inquire whether a dictionary type of list could be used to assist in the encoding into sentences. It is certainly clear that it would be impossible to list all of the sentences of the language in a dictionary. In fact, an attempt to list all two-word sequences would require a dictionary of impractical size. The length of the list required to accommodate all structures of a code depends upon the redundancy of the structures, but more important, upon the size of the signaling alphabet and the length of the sequences. The use of words as a signaling alphabet and the use of sequences of sentence length is completely out of question because of the practical impossibility of listing and storing enough sentences.

In order to reduce the signaling alphabet, the concept of part of speech is introduced. Larger structures are stated in terms of sequence of parts of speech instead of sequences of words. By the introduction of the concept of part of speech, we have factored the message into two parts. First of all, there is a sentence composed of a sequence of parts of speech, and the encoder has the opportunity of choice from among the various allowed sequences. Second, there is a further opportunity for choice from among the words that have the privilege of occurrence¹⁰ for each part of speech. In language, these two possibilities for choice correspond to structural meaning and lexical meaning. As an illustration of structural meaning, take the sentence, "The man had painted the house."

A German sentence with approximately the same meaning as the one above, translated on a word-for-word basis, would be, "The man had the house painted." Here the words are the same, but the structural meaning is different.

As an example of the economy introduced by the concept of part of speech, consider the Markov source (See Fig. 3.) which will generate over 10^{21} English sentences using a vocabulary of about 35 words. By the use of the concept of part of speech, whole lists of words are considered as equivalent so that with the 10 parts of speech there is only a small number of sentence types. It is estimated that there are millions of possible sentence types of which this diagram represents only a few. The structural meaning is indicated by the sentence type or the choice of path through the diagram, the lexical meanings are indicated by the further choice of the individual words from each list.

The introduction of part of speech and the factoring of the message into a lexical and a structural part has reduced the total number of the possible representations of sentences. The number of different structures, however, is still too large to list in a dictionary. The further step that we propose to take is to take advantage of regularities in the sentence types. For example, the first three states in the diagram (Fig. 3) and their connecting lines may be found intact in many different sentence types and often more than once in a given sentence type. Just as we have grouped several words together to make a part of speech, we may group several paths together to form a phrase. If this program is carried out in its full elaboration, we are left with a number of intermediate levels of structure between the word and the sentence, such as various types of phrases and clauses. The levels are to be chosen in such a way that the total number of listed structures is reduced to a number that can be handled in a machine memory. Preliminary work seems to show that this can be achieved if the parts of speech number in the hundreds.

As an illustration of the use of an analogous level structure in coding, we can turn to the error-proof codes of Elias¹¹. In these codes, "words" are formed according to some error-correcting code, such as one of those already mentioned, in which there are message digits and check digits. After a sequence of words has been sent, a phrase is made by adding a series

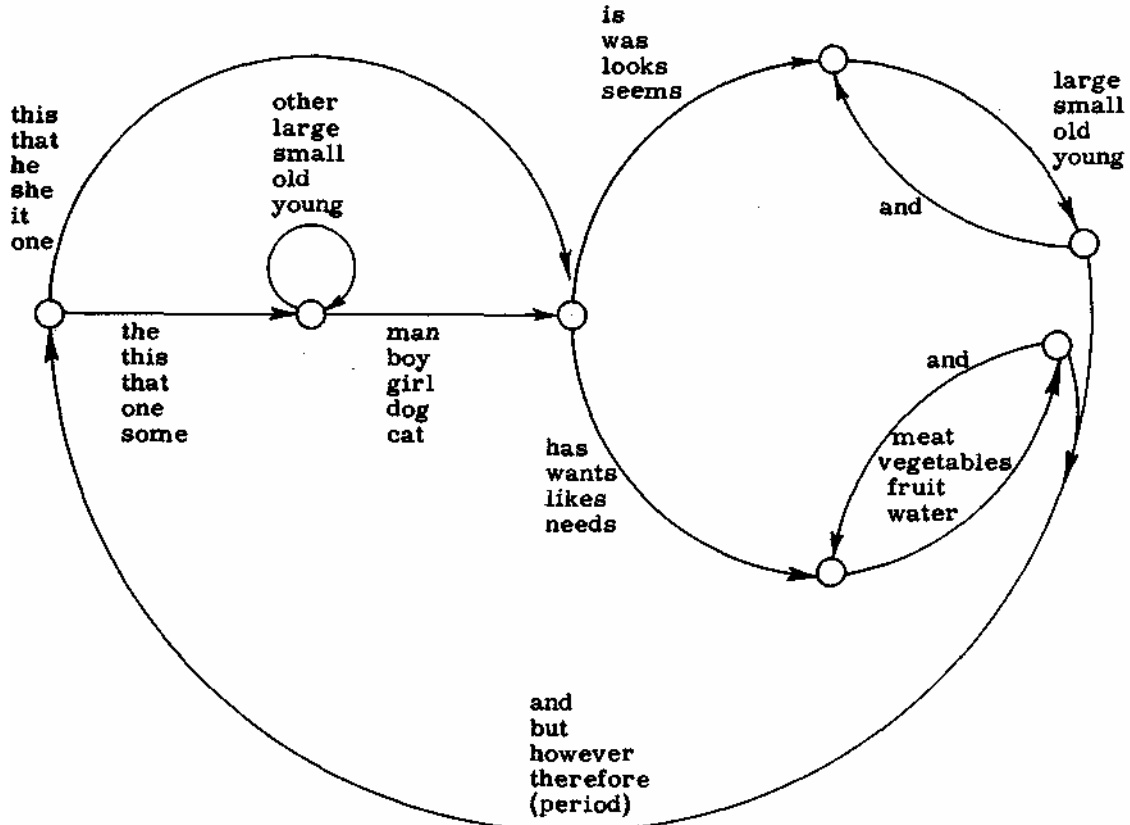


Fig. 3

of check words so that the whole structure has error-correcting properties on the phrase level as well as on the word level. The process is iterated as often as desired.

A somewhat closer analogy to language could be constructed by dividing the words into parts of speech (indicated, for instance, by the first digit) so that we would have two parts of speech). A sentence of seven words in this code is represented by the seven rows of the diagram (Fig. 4). The structural meaning

A	I	I	I	C	C	C
A	I	I	I	C	C	C
A	I	I	I	C	C	C
A	I	I	I	C	C	C
B	I	I	I	C	C	C
B	I	I	I	C	C	C
B	I	I	I	C	C	C

Fig. 4

is indicated by the binary digits marked A, and these are checked by check digits marked B. The lexical meanings are indicated by the rows of III. In each word, AIII or BIII is

checked by the digits C. In this code, the parts of speech are clearly and explicitly marked in the absence of noise by certain features (the first digit) in each word; in language, parts of speech are not always very clearly marked by grammatical affixes or the like. In language, there is no explicit separation into message symbols and symbols furnished by the constraints of the code, but our assumption that each sentence can be translated into another language leads us to look for an implicit separation.

Our rules of language from the point of view of the encoder, then, are somewhat as follows. Select a sentence from among the sequences of clause types. For each clause type, select a clause from among the allowed sequences of phrase types. For each phrase, select a sequence of parts of speech. For each part of speech, select a word. In the translation process, the information required for the selections at each stage must be obtained from the decoder and may be called the "message" represented in the transition language.

Structure of Language from the Point of View of the Decoder

So far, the structure of language has been looked at from the point of view of the encoder which encodes in a given output language the "message" provided for it by the decoder. The rules for decoding language into some representation of the "message" are not just the reverse of the rules for encoding. If they were, mechanical translation would be much easier to accomplish than it appears to be. The difference between the point of view of the decoder and the encoder is just the difference between analysis and synthesis. The difference is illustrated in error-correcting codes that are easy to encode according to rules, but for which no rules are known for decoding in the presence of noise, although the message can be recovered by the use of a code book. In language, the difficulties in decoding are not the result of noise; they are the result of certain characteristics of the encoding scheme.

Decoding would be very simple with the error-correcting code using two parts of speech (Fig. 4). Decoding would be simple and direct because the part of speech of each word is clearly marked by its first digit. This is true to a certain extent in languages that have inflectional endings and grammatical affixes; more so in some languages than in others. Much attention has been paid to these affixes for purposes of mechanical translation. But the fact remains that even in the most highly inflected languages, the parts of speech are imperfectly indicated by affixes on the words. The problem is even worse than that: a given word form may belong to more than one part of speech, and there is no way at all to tell which part of speech it is representing in a certain sentence by looking at the word itself. The context, or the rest of the sentence must be examined. The lists of words that the encoder uses for each part of speech overlap, so that a given word may appear on several lists. In Fig. 3 it can be seen that several of the words appear in more than one list. The proper translation of these words into a language other than English requires a knowledge of the list from which the word was chosen. The decoder has this problem of deducing from which list the word was chosen. The statement that a word may belong to several parts of speech is just another way of saying that it may have several meanings. The concept of part of speech may

be extended to include not only the usual grammatical distinctions, but in addition the distinctions that usually would be called multiple meanings.

Probably all languages exhibit the phenomena of multiple meaning, and one word making shift for more than one part of speech. It is interesting to speculate as to whether there is any utility to this phenomena, or whether it is just excess baggage, a human failing, another way in which our language does not come up to ideal. One word — one meaning would presumably make our language more precise and would eliminate the basis for many pointless arguments and much genuine misunderstanding. It has been proposed that language be changed to approach the ideal of one word — one meaning so that mechanical translation would be easier¹². Some of the advantages accruing from the phenomena of multiple meaning might be as follows: There is an economy of the vocabulary because part of the burden of carrying meaning is transferred to the word sequence. The number of different structures available in a code goes as V^n , where V is the vocabulary size and n is the length of the sequences. In order to take advantage of the larger number of structures available, the words must acquire multiple meanings. There is the introduction of the possibility of the metaphoric extension of the meaning of words so that old words can be used for new concepts. There is the possibility of using a near synonym if a word with the exact meaning is not at hand, and of modifying the meaning of the near synonym to that intended by putting it in an appropriate context.

Since the lists of words for the different parts of speech used by the encoder overlap, there is the possibility that the same sequence of words may result from different intended structural meanings. In fact, this sometimes happens when the encoder is not careful, and we have a case of ambiguity. Sometimes the choice of an ambiguous sequence is intentional, and we have a pun. Puns, in general, cannot be translated, and we have to assume that unintentional ambiguity is at a minimum in the carefully written material that we want to translate.

The task of the decoder in a translation process is to furnish the information required by the encoder so that it can make the appropriate selections on each level of structure. This information is implicit in the incoming sequence

of words and must be made explicit. The decoder is given only the words of the incoming text and their arrangement into sentences. It must reconstruct the assignment of the words to the parts of speech intended by the encoder, and must make the structural meaning explicit so that it can be translated. The decoder must resolve the problems of multiple meaning of words or structures in case these meanings are expressed in several ways in the other language. The decoder has available two things: the words, and the context surrounding each of the words. The appropriate starting point for describing the structure of language from the point of view of the decoder is to classify the words of the language and the contexts of the language. The classification proceeds on the assumption that there is no ambiguity, that the assignment of words to parts of speech can be done by the decoder either by examining the form of the words themselves or by examining the context.

The classification of the words must be a unique one. Each word must be assigned to one and only one class. These we shall call word classes. In order to set up word classes, we classify together all word forms that are mutually substitutable in all sentences and behave similarly in translation. In practice, one of the difficulties of making such a classification is the problem of how detailed the classification should be. Certain criteria of usage must be ignored or in the end each word class will have only one word in it. As examples of the sort of classification that is intended, "a" and "the" would be assigned to different classes because "a*" cannot be used with plural nouns. "To" and "from" would be assigned to different word classes because "to" is a marker of the infinitive. "Man" and "boy" would be assigned to different word classes because you can man a boat. But "exact" and "correct" would not be separated merely because one can exact a promise but correct an impression. Preliminary experimentation has indicated that the number of word classes needed for translating the structural meaning is of the order of many hundreds.

The classification of contexts is very closely connected with the setting up of word classes. A sentence can be considered as a sequence of positions. Each position is filled by a word and surrounded by a context. Since we have classified words into word classes, each

position in the sentence has associated with it a word class which can be determined uniquely by looking the word up in a special dictionary. The number of sentence length sequences of word classes is much fewer than the number of sentences. All sentences that have the same sequence of word classes are considered equivalent. The context of a given position in a sentence can be represented by the sequence of word classes preceding the position and the sequence of word classes following the position, but all within one sentence length. It is these contexts that we propose to classify. We classify together all contexts that allow the substitution of words from the same set of word classes. We thus have set up both word classes and context classes.

The relationship between the word classes and the context classes can be illustrated by a very large matrix. The columns of the matrix represent all of the word positions in any finite sample of the language. The rows of the matrix represent different word forms in the vocabulary of the language. Each square in the matrix is marked with an X if the word corresponding to that row will fit into the context surrounding the position corresponding to that column. All words that have identical rows of X's belong to the same word class. All contexts that have identical columns of X's belong to the same context class.

The word classes and the context classes can be set up in such a way that the sentence sequence of context classes contains just the information that we require for specifying the original parts of speech — and thus the structural meanings — as well as the information that we require for resolving many of the multiple meanings of the words and of the larger structures.

The structure of language from the point of view of the decoder is as follows. Words are listed in a dictionary from which we can obtain for each its assignment to a word class. Sequences of word classes are also listed in the dictionary, together with their designations in terms of phrase types. Sequences of these phrase types are also listed in the dictionary, and so on, until we have sentence types. The procedure for the decoder is to look up in the dictionary the longest sequences that it can find listed, proceeding from word class sequences to phrase sequences, to clause sequences and so on. At each look-up step, the dictionary gives explicit

expressions that lead in the end to a discovery of the context classes of each position. From this we obtain, for each word, its original assignment to a part of speech, and the structural meaning. Thus we have the "message" or explicit directions for use in the encoder.

Conclusion

The mechanical translation of languages on a sentence-for-sentence basis is conceived of as a two-step process. First, the incoming text is decoded by means of a decoder working with the constraints of the input language expressed in dictionary form and based on word classes and context classes. The result of the decoding operation is a representation of the "message," which is just the directions that the encoder needs to re-encode into the output language by using the constraints of the output language expressed in dictionary form and based on parts of speech. An assessment of the worth or the fidelity of the resulting translations must await completion of the detailed work required to set up the dictionaries and to work out the system in all detail. It is certain that the resulting translations will be better than any word-for-word translations.

Acknowledgment

The author is deeply appreciative of the opportunity that he has had for discussing these matters with his colleagues at the Research Laboratory of Electronics, Massachusetts Institute of Technology. He is particularly indebted to R. F. Fano, P. Elias, F. Lukoff, and N. Chomsky for their valuable suggestions and comments.

References

1 An earlier version of some of the ideas contained in this paper can be found in Chapter 14 of reference 2.

2 Machine Translation of Languages, edited by W. N. Locke and A. D. Booth, The Technology Press of M.I.T. and John Wiley and Sons, Inc., New York; Chapman and Hall, Ltd., London (1955).

3 See various issues of Mechanical Translation, a journal published at Room 14N-307, Massachusetts Institute of Technology, Cambridge 39, Mass., U.S.A.

4 Page 45 of reference 2.

5 R. W. Hamming, "Error detecting and error correcting codes," Bell System Tech. J. 31, 504-522 (1952).

6 D. E. Muller, "Metric Properties of Boolean Algebra and their Application to Switching Circuits," Report No. 46, Digital Computer Laboratory, University of Illinois (April 1953).

7 I. S. Reed. "A class of multiple error-correcting codes and the decoding scheme," Trans. I.R.E. (PGIT) 4. 38-49 (1954).

8 Y. Bar-Hillel, "The present state of research on mechanical translation," American Documentation. 2, 229-237 (1951).

9 E. Reifler, "Mechanical determination of the constituents of German substantive compounds," Mechanical Translation. II, No. 1 (July. 1955).

10 L. Bloomfield, Language, Henry Holt and Company, Inc., New York (1933).

11 P. Elias, "Error-free coding," Trans. I.R.E. (PGIT) 4, 30-37 (1954).

12 Chapter 10 of reference 2.