LINGUISTICS AND COMMUNICATION THEORY

Norbert Wiener refuses to admit "any fundamental opposition between the problems of our engineers in measuring communication and the problems of our philologists".[1] There appear indeed striking coincidences and convergences between the latest stages of linguistic analysis and the approach to language in the mathematical theory of communication. Since each of these two disciplines is concerned, although in different and quite autonomous ways, with the same domain of verbal communication, a close contact between them has proved to be of mutual use and undoubtedly will become ever more beneficial.

The stream of oral speech, physically continuous, originally confronted the mathematical theory of communication with a situation "considerably more involved" than in the case of a finite set of discrete constituents, as presented by written speech.[2] Linguistic analysis, however, came to resolve oral speech into a finite series of elementary informational units. These ultimate discrete units, the so-called "distinctive features", are aligned into simultaneous bundles termed "phonemes", which in turn are concatenated into sequences. Thus form in language has a manifestly granular structure and is subject to a quantal description.

The primary aim of information theory, as stated for instance by D. M. MacKay, is "to isolate from their particular contexts those abstract features of representations which can remain invariant under reformulation".[3] The linguistic analogue to this problem is the phonemic search for relational invariants. The diverse possibilities for measurement of the amount of phonemic information which are foreseen by the communication engineers (distinguishing between "structural" and "metrical"

---

[1] *Journal of the Acoustical Society of America*, vol. 22 (1950), p. 697.
[2] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (Urbana, The University of Illinois Press, 1949), pp. 74 ff., 112 f.
[3] *Cybernetics: Transactions of the Eighth Conference* (New York, Josiah Macy Jr. Foundation, 1952), p. 224.

information content) may give both to synchronic and to historical linguistics precious data, particularly important for the typology of languages both in a purely phonological aspect and in the intersection of phonology with the lexicogrammatical level.

The dichotomous principle underlying the whole system of distinctive features in language has gradually been disclosed by linguistics and has found corroboration in the binary digits (or to use the popular portmanteau, *bits*) employed as a unit of measurement by the communication engineers. When they define the selective information of a message as the minimum number of binary decisions which enable the receiver to reconstruct what he needs to elicit from the message on the basis of the daa already available to him,[4] this realistic formula is perfectly applicable to the role of distinctive features in verbal communication. As soon as "the way of recognizing universals through their invariants" had been attempted and an overall classification of distinctive features had been outlined according to these principles, the problem of translating the criteria proposed by linguists "into mathematical and instrumental language" was raised in D. Gabor's lectures on communication theory.[5] Recently there has appeared an instructive study by G. Ungeheuer offering a tentative mathematical interpretation of the distinctive features in their binary patterning.[6]

The notion of "redundancy", taken over by communication theory from the rhetorical branch of linguistics, acquired an important place in the development of this theory, has been challengingly redefined as "one minus the relative entropy", and under this new aspect has reentered present-day linguistics as one of its crucial topics. The necessity of a strict distinction between different types of redundancy is now realized in the theory of communication as well as in linguistics, where the concept of redundancy encompasses on the one hand pleonastic means as opposed to explicit conciseness (*brevitas* in the traditional nomenclature of rhetoric) and on the other hand explicitness in contradistinction to ellipsis. On the phonological level, linguists have been accustomed to delimit phonemic, distinctive units from contextual, combinatory, allophonic variants, but the treatment of such interconnected problems as redundancy, predictability, and conditional probabilities in communication theory furthered a clarification of the relationship between the two basic linguis-

---

[4] *Communication Theory*, ed. by W. Jackson (New York, Academic Press, 1953), p. 2.
[5] *Lectures on Communication Theory* (Massachusetts Institute of Technology, Cambridge, Mass., 1951), p. 82.
[6] *Studia Linguistica*, vol. 13 (1959), pp. 69–97.

tic classes of sound-properties – the distinctive features and the redundant features.

A phonemic analysis, when consistently purposing the elimination of redundancies, necessarily provides an optimal and unambiguous solution. The superstitious *belief* of some theoreticians unconversant with linguistics that "there remain no good reasons for the distinction between distinctive and redundant among the features"[7] is patently contradicted by innumerable linguistic data. If, for example, in Russian the difference between advanced vowels and their retracted counterparts is always accompanied by the difference between preceding consonants, which are palatalized before the advanced vowels and devoid of palatalization before the retracted vowels, and if on the other hand the difference between palatalized and non-palatalized consonants is not confined to a vocalic neighborhood, the linguist is obliged to conclude that in Russian the difference between the presence and lack of consonantal palatalization is a distinctive feature, while the difference between the advanced and retracted vowels appears as merely redundant. Distinctiveness and redundancy, far from being arbitrary assumptions of the investigator, are objectively present and delimited in language.

The prejudice treating the redundant features as irrelevant and distinctive features as the only relevant ones is vanishing from linguistics, and it is again communication theory, particularly its treatment of transitional probabilities, which helps linguists to overcome their biased attitude toward redundant and distinctive features as irrelevant and relevant respectively.

Preconceived possiblitities, according to MacKay, "is the key phrase in communication theory", and a similar claim comes from linguistics. In neither discipline has there been any doubt about the fundamental role of selective operations in verbal activities. The engineer assumes a "filing system" of prefabricated possibilities more or less common to the sender and receiver of a verbal message, and Saussurian linguistics speaks correspondingly about *langue*, which makes possible an exchange of *parole* between interlocutors. Such an "*ensemble* of possibilities already foreseen and provided for"[8] implies a code, conceived by communication theory as "an agreed transformation – usually one-to-one and reversible"[9] – by which one set of informational units is converted into another set, for

---

[7] *Word*, vol. 13 (1957), p. 328.

[8] *Cybernetics: Transactions of the Eighth Conference* (New York, The Technology Press of Massachusetts Institute of Technology, 1952), p. 183.

[9] C. Cherry, *On Human Communication* (New York-London, 1957), p. 7.

instance, a grammatical unit into a phonemic sequence and vice versa. The code matches the *signans* with its *signatum* and the *signatum* with its *signans*. Today, with respect to the treatment of coding problems in communication theory, the Saussurian dichotomy *langue/parole* can be restated much more precisely and acquires a new operational value. Conversely, in modern linguistics communication theory may find illuminating information about the stratified structure of the intricate linguistic code in its various aspects.

Although the framework of the linguistic code has been adequately outlined in linguistics, it is still frequently overlooked that the finite ensemble of "standard representations" is limited to lexical symbols, their grammatical and phonological constituents, and the grammatical and phonological rules of combination. Only this portion of communication may be defined as a mere "activity of replicating representations". On the other hand, it is still opportune to recall that the code is not confined to what communication engineers call "the bare intelligence content" of speech, but that likewise the stylistic stratification of the lexical symbols and the allegedly "free" variation, both in their constitution and in their combination rules, are "foreseen and provided for" by the code.

In his program for the future science of signs (semiotics) Charles Peirce stated: "A Legisign is a law that is a Sign. This law is usually established by men. Every conventional sign is a legisign."[10] Verbal symbols are cited as a salient example of legisigns. Interlocutors belonging to one given speech community may be defined as actual users of one and the same linguistic code encompassing the same legisigns. A common code is their communication tool, which actually underlies and makes possible the exchange of messages. Here is the essential difference between linguistics and the physical sciences, and this difference has been distinctly and repeatedly singled out in the theory of communication, especially by its English school, which insists on a clear-cut line of demarcation between the theory of communication and of information. Nevertheless, this delimitation, strange as it seems, is sometimes disregarded by linguists. "The stimuli received from Nature," as Colin Cherry wisely stresses, "are not pictures of reality but are the evidence from which we build our personal models."[11] While the physicist creates his theoretical construct, imposing his own hypothetical system of new symbols upon

---

[10] *Collected Papers*, vol. 2 (Cambridge, Mass., Harvard University Press, 1932), p. 142f.
[11] *Op. cit.*, p. 62. Cf. W. Meyer-Eppler, *Grundlagen und Anwendungen der Informationstheorie* (Berlin-Göttingen-Heidelberg, Springer-Verlag, 1959), p. 250ff.

the extracted *indices*, the linguist only recodes, translates into symbols of a metalanguage those extant *symbols* which are used in the language of the given speech community.

The constituents of the code, for instance, the distinctive features, literally occur and really function in speech communication. Both for the receiver and for the transmitter, as R. M. Fano points out, the operation of selection forms the basis of "information-conveying processes".[12] The set of yes-or-no choices underlying any bundle of these discrete features is not an arbitrary concoction of the linguist, but is actually made by the addressee of the message, insofar as the need for their recognition is not cancelled by the prompting of the verbal or non-verbalized context.

Both on the grammatical and on the phonological level, not only the addressee in decoding the message, but also the encoder may practice ellipsis; in particular, the encoder omits some of the features, or even some of their bundles and sequences. But ellipsis, too, is governed by codified rules. Language is never monolithic; its overall code includes a set of subcodes, and such questions as that of the rules of transformation of the optimal, explicit kernel code into the various degrees of elliptic subcodes and their comparison as to the amount of information requires both a linguistic and an engineering examination. The *convertible code* of language, with all its fluctuations from subcode to subcode and with all the current progressing changes which this code is undergoing, is to be jointly and comprehensively described by the means of linguistics and communication theory. An insight into the dynamic synchrony of language, involving the space-time coordinates, must replace the traditional pattern of arbitrarily restricted *static* descriptions.

The linguistic observer who possesses or acquires a command of the language he is observing is or gradually becomes a potential or actual partner in the exchange of verbal messages among the members of the speech community, a passive or even active fellow member of that community. The communication engineer is right when defending against "some philologists" the absolutely dominant "need to bring the Observer onto the scene" and when holding with Cherry that "the participant-observer's description will be the more complete".[13] The antipode to the participant, the most detached and external onlooker, acts as a cryptanalyst, who is a recipient of messages without being their addressee

---

[12] *The Transmission of Information* (Massachusetts Institute of Technology, Research Laboratory of Electronics, Technical Report No. 65) (1949), p. 3f.
[13] *For Roman Jakobson* (The Hague, Mouton & Co., 1956), p. 61f.

and without knowledge of their code.[14] He attempts to break the code through a scrutiny of the messages. As far as possible, this level of linguistic investigation must be merely a preliminary stage toward an internal approach to the language studied, when the observer becomes adjusted to the native speakers and decodes messages in their mother-tongue through the medium of its code.

As long as the investigator knows no *signatum* of a given language and has access to nothing but *signans* he willy-nilly has to strain his detective capacities and obtain any possible information about the structure of this language from the external evidence. The present state of Etruscology is a good example of such a technique. But if the linguist is familiar with the code and has mastered the conventional transformations by which a set of *signantia* is converted into a set of *signata*, then it becomes super-fluous for him to play Sherlock Holmes, unless he aspires to finding out how wide and reliable data could be obtained by such a mock scrutiny. It is difficult, however, to simulate ignorance of a familiar code: smuggled-in meanings distort one's allegedly cryptanalytic approach.

Obviously "the inseparability of objective content and observing sub-ject", singled out by Niels Bohr as a premise of all well-defined knowledge,[15] must be definitely taken into account also in linguistics, and the position of the observer in relation to the language observed and described must be exactly identified. First, as formulated by Jurgen Ruesch, the infor-mation an observer can collect depends upon his location within or out-side the system.[16] Furthermore, if the observer is located *within* the communication system, language presents two considerably different aspects when seen from the two ends of the communication channel. Roughly, the encoding process goes from meaning to sound and from the lexicogrammatical to the phonological level, whereas the decoding process displays the opposite direction – from sound to meaning and from features to symbols. While a set (*Einstellung*) toward immediate constituents takes precedence in speech production, for speech perception the message is *first* a stochastic process. The probabilistic aspect of speech finds conspicuous expression in the approach of the listener to homonyms, whereas for the speaker homonymy does not exist. When saying /sʌn/, he knows beforehand whether "sun" or "son" is meant, while the listener depends on the conditional probabilities of the context.[17] For the receiver,

[14] Cf. R. Jakobson and M. Halle, *Fundamentals of language* (The Hague, Mouton & Co., 1956), pp. 17–19.
[15] *Atomic Physics and Human Knowledge* (New York, John Wiley & Sons, 1958), p. 30.
[16] *Toward a Unified Theory of Human Behavior*, ed. by R. R. Grinker (New York, Basic Books, 1956), p. 54.
[17] See *International Journal of Slavic Linguistics and Poetics*, vol. 1/2 (1959), p. 286f.

the message presents many ambiguities which were unequivocal for the sender. The ambiguities of pun and poetry utilize this input property for the output.

No doubt there is a feedback between speaking and hearing, but the hierarchy of the two processes is opposite for the encoder and decoder. These two distinct aspects of language are irreducible to each other; both are equally essential and must be regarded as *complementary* in Niels Bohr's sense of this word. The relative autonomy of the input pattern is documented by the widespread temporal priority of a passive acquisition of a language both by infants and adults. L. Ščerba's request for a delimitation and elaboration of two grammars – "active" and "passive" – recently reemphasized by young Russian scholars is equally important for linguistic theory, for language teaching, and for applied linguistics.[18]

When a linguist deals with one of the two aspects of language *à al Jourdain*, namely without realizing whether his observations are devoted to output or to input, that is still less dangerous than the arbitrary compromises frequently made between the output and input analyses – for example, an output grammar discussing general operations without appeal to meaning, despite the necessary priority of meaning for the encoder. At present, linguistics is receiving particularly valuable suggestions from communication theory for the somewhat neglected study of the verbal input.

MacKay warns against the confusion between the exchange of verbal messages and the extraction of information from the physical world, both abusively unified under the label "communication"; this word has for him an inevitably *anthropomorphic* connotation "which bedevils discussion".[19] There is a similar danger when interpreting human inter-communication in terms of physical information. Attempts to construct a model of language without any relation either to the speaker or to the hearer, and thus to hypostasize a code detached from actual communication, threaten to make a scholastic fiction out of language.

Beside encoding and decoding, also the procedure of recoding, code switching, briefly, the various facets of translation, is becoming one of the focal concerns both of linguistics and of communication theory, here and in Western and Eastern Europe. Only now do such fascinating problems as those of ways and degrees of mutual understanding among speakers of

[18]   See I. Revzin, *Tezisy konferencii po mašinnomu perevodu* (Moscow, Pervyj Moskov. Gos. Ped. Inst. Inostrannyx Jazykov, 1958), pp. 23–25.
[19]   *Cybernetics: Transactions of the Eighth Conference* (New York, 1952), p. 221.

such closely cognate languages, as, for instance, Danish, Norwegian and Swedish, begin to attract the attention of linguists[20] and promise to give lucid insight into the phenomenon known in communication theory under the label "semantic noise" and into the theoretically and pedagogically important problem of overcoming it.

Incidentally, both linguistics and communication theory for a certain period were tempted to treat any concern with meaning as a kind of semantic noise and to exclude semantics from the study of verbal messages. At present, however, linguists display a tendency to reintroduce meaning while utilizing the very instructive experience of this temporary ostracism. Also in communication theory a similar trend may be observed. According to Weaver, the analysis of communication "has so penetratingly cleared the air that one is now, perhaps for the first time, ready for a real theory of meaning", and especially for handling "one of the most significant but difficult aspects of meaning, namely the influence of context".[21] Linguists are gradually finding the way of tackling meaning and in particular the relation between general and contextual meaning as an intrinsic linguistic topic, distinctly separate from the ontological problems of reference.

Communication theory, after having mastered the level of phonemic information, may approach the task of measuring the amount of grammatical information, since the system of grammatical, particularly morphological categories, like the system of distincitive features, is ostensibly based on a scale of binary oppositions. Thus, for instance, 9 binary choices underlie over 100 simple and compound conjugational forms of an English verb which appear, for example, in combination with the pronoun *I*.[22] The amount of grammatical information carried by the English verb can subsequently be confronted with the corresponding data on the English noun or on the verb and noun of various languages; the relation between the morphological and syntactic information in English has to be compared with the equivalent relation in other languages, and all these comparative data will present important auxiliary material for a

[20] See particularly E. Haugen, *Nordisk Tidskr.*, vol. 29 (1953), pp. 225–249.

[21] Shannon and Weaver, *op. cit.*, p. 116. Cf. D. M. MacKay, "The place of 'meaning' in the theory of information", *Information Theory*, ed. by C. Cherry (New York, Basic Books, 1956).

[22] 1. Preterit (*vs.* non-preterit), 2. perfect, 3. progressive, 4. expective, 5. morally determined, 6. contingent, 7. potential, 8. assertorial, 9. passive. Cf. R. Jakobson, *American Anthropologist*, vol. 61, no. 5, Part 2 (1959), pp. 139–141 [above, pp. 489ff.]; W. F. Twaddell, *The English Verb Auxiliaries* (Providence, Brown University Press, 1960).

linguistic typology of languages and for the inquiry into linguistic universals.

The amount of grammatical information which is potentially contained in the paradigms of a given language (statistics of the code) must be further confronted with a similar amount in the tokens, in the actual occurrences of the various grammatical forms within a corpus of messages. Any attempt to ignore this duality and to confine linguistic analysis and calculation only to the code or only to the corpus impoverishes the research. The crucial question of the relationship between the patterning of the constituents of the verbal code and their relative frequency both in the code and in its use cannot be passed over.

The semiotic definition of a symbol's meaning as its translation into other symbols finds an effectual application in the linguistic testing of intra- and interlingual translation, and this approach to semantic information concurs with Shannon's proposal to define information as "that which is invariant under all reversible encoding or translating operations", briefly, as "the equivalence class of all such translations".[23]

When dealing with meanings, whether grammatical or lexical, we must be careful not to misuse the polar notions – "regularity" and "deviation". The idea of deviation frequently arises from a disregard for the stratified, hierarchic structure of language. There is, however, a substantial difference between secondariness and deviation. We are not justified in envisaging as deviant either Kuryłowicz's "syntactic derivation" with regard to the "primary function",[24] or Chomsky's "transforms" *versus* "kernels",[25] or Bloomfield's "marginal" ("transferred") meanings in relation to the "central" meaning of the word.[26] Metaphoric creations are not deviations but regular processes of certain stylistic varieties, which are subcodes of an overall code, and within such a subcode there is nothing deviant in Marvell's figurative assignment of a concrete epithet to an abstract noun (properly a hypallage) – "a green Thought in a green shade" – or in Shakespeare's metaphoric transposition of an inanimate noun into the feminine class – "the morning opes her golden gates" – or in the metonymic use of "sorrow" instead of "sorrowful while", which Putnam's paper excerpts from Dylan Thomas ("A grief ago I saw him there").[27] In contradistinction to such agrammatical construc-

---

[23] *Cybernetics: Transactions of the Seventh Conference* (New York, 1951), p. 157.

[24] *Bulletin de la Société de Linguistique de Paris*, no. 110 (1936), pp. 79–92.

[25] *Syntactic Structures* (The Hague, Mouton & Co., 1957).

[26] *Language* (New York, Henry Holt & Co., 1933), p. 149.

[27] *Proceedings of Symposia in Applied Mathematics*, XII (1961).

tions as "girls sleeps", the quoted phrases are meaningful, and any meaningful sentence can be submitted to a truth test, exactly in the same way as the statement, "Peter is an old fox" could lead to a reply, "It's not true; Peter is not a fox but a swine; but John is a fox." Incidentally, neither ellipsis nor reticence or anacoluthon can be considered as deviant structures; they, and the slurred style of speech, a brachylogical subcode to which they belong, are merely lawful derivations from the kernel forms embedded in the explicit standard. Once again, this "code variability", which clarifies why the standard is not actualized in some overt behavior, has been overlooked more by linguists than by the less "biased" communication engineers.

To sum up, there exists a wide range of questions calling for cooperation of the two different and independent disciplines we are discussing. The first steps in that direction were actually lucky. May I finish by quoting an example of the longest and, until recently, perhaps the most spectacular tie between linguistics, in particular the study of poetic language, on the one hand, and the mathematical analysis of stochastic processes on the other. The Russian school of metrics owes some of its internationally echoed achievements to the fact that some forty years ago such students as B. Tomaševskij, expert both in mathematics and in philology, skillfully used Markov chains for the statistical investigation of verse; these data, supplemented by a linguistic analysis of the verse structure, gave in the early twenties a theory of verse based on the calculus of its conditional probabilities and of the tensions between anticipation and unexpectedness as the measurable rhythmical values, and the computation of these tensions, which we have labeled "frustrated expectations", gave surprising clues for descriptive, historical, comparative, and general metrics on a scientific basis.[28]

I am convinced that methods newly developed in structural linguistics and in communication theory, when applied to verse analysis, and to many other provinces of language, are capable of opening up wide perspectives for further coordinated efforts of both disciplines. Let us anticipate that our expectations will not be frustrated.[29]

[28]   Cf. B. Tomaševskij, *O stixe* (Leningrad, 1929); R. Jakobson, *O češskom stixe* (Berlin-Moscow, 1923), and "Linguistics and poetics", *Style in Language* (New York, The Technology Press of Massachusetts Institute of Technology, 1960).
[29]   I should like to dedicate this paper to the memory of the engineer O. A. Jakobson.

ROMAN
JAKOBSON

SELECTED
WRITINGS

II

Word and Language

D.   PHILOLOGICAL GLEANINGS

E.   INTERDISCIPLINARY OUTLOOKS