

ON THE IDENTIFICATION OF PHONEMIC ENTITIES*

ROMAN JAKOBSON

As phonemes are linguistic elements, it follows that no phoneme can be correctly defined except by linguistic criteria, i.e., by means of its function in the language. No extra-lingual criteria can be relevant, i.e., neither physical nor physiological nor psychological criteria.

This motto, extracted from Louis Hjelmslev's address to the Second International Congress of Phonetic Sciences (London, 1935),¹ is one of the most stimulating of the many ideas which he contributed to the development of structural linguistics. May I now briefly trace here the corollaries of this sagacious methodological requirement as I, for my part, conceive them.

"Both phonematic, grammatical and lexical elements," the cited paper insists, "are at the same time inner and outer phenomena." As a matter of fact, this twofold nature of any phonemic entity, uniting it with all the superordinated linguistic constituents, was ascertained at the very outset of phonemic inquiry. Beginning in the early eighties, Baudouin de Courtenay repeatedly stated that from the semantic point of view "the utterance breaks up into sentences, sentences into significative words, words into morphological components or morphemes and morphemes into phonemes." For "a morpheme is divisible only into components which are of the same nature as itself: they, too, must be significative." Thus, a dissociation of the morpheme into physical or physiological elements, i.e., into sounds, would be, according to Baudouin de

Courtenay, "an unjustified and paralogical jump in division."

Likewise for Ferdinand de Saussure any linguistic constituent is necessarily a "two-fold entity" ("entité à deux faces"). "Une unité matérielle n'existe que par le sens, la fonction dont elle est revêtue; ce principe est particulièrement important pour la connaissance des unités restreintes, parce qu'on est tenté de croire qu'elles existent en vertu de leur pure matérialité. . . . Inversement . . . un sens, une fonction n'existent que par le support de quelque forme matérielle."

The lesson taught by both great discoverers of the primary concepts of structural linguistics is quite unequivocal. But when the time came for linguistics to utilize these fundamentals and to build up a total conception of language far removed from tradition both in theory and in descriptive technique, it proved much more difficult to heed the cited methodological warnings than merely to admit their validity.

The 1930s witnessed an impetuous and promising international development of phonemic studies, and Hjelmslev's polemic fervor in focusing attention on and vindicating the strictly *linguistic criteria* was most appropriate and expedient.

However various were the definitions of the phoneme offered by different scholars and schools, all of these formulations aimed at essentially one and the same thing, and, in broad outline, the practical task of enumerating the stock of phonemes for any given language found its approximate solution. The difficulties began with the attempts at the in-

¹The *Proceedings of the Second International Congress of Phonetic Sciences* (Cambridge, 1936), p. 49.

*This article was first published in the *Travaux du Cercle Linguistique de Copenhague* 5.205-213 (1949). Reprinted here from Roman Jakobson, *Selected Writings I* (Mouton, 1962), pp. 418-425, by permission.

ward characterization of any of these phonemes, and two diverse attitudes are most indicative for the whole initial period of phonemic investigation. Either the student limited himself to registering the number of distinct phonemes in the given language and to symbolizing graphically the mere fact of this distinctness, or he made the obviously paralogical jump from a purely linguistic analysis to the raw sound-matter: he defined, e.g., the French phoneme *g* as a postalatal stop, *n* as an antepalatal nasal, and *ʒ* as a palato-alveolar spirant, without asking himself whether, *ceteris paribus*, the difference between the articulation toward these three regions of the palate could be of some autonomous relevance for the phonemic distinction of French consonants.

It seems to me that the consistent application of linguistic criteria as required by Hjelmslev does not allow the discontinuation of the phonemic analysis either through a kind of agnosticism toward the proper essence of the phonemes or through a substitution of sound-physiology and physics for a thoroughly linguistic procedure.

Overcoming the one-track mind of the neogrammarian bias, F. de Saussure pointed out that beside the *axis of successiveness*, linguistics, as any science dealing with *values* does, must also tackle the other coordinate—the *axis of simultaneity* “concerning relations between coexistent things”; whereas for the traditional approach there was no science of language outside questions of succession. The neogrammarians were entirely taken up with the axis of successiveness, as if it were possible to comprehend the sequence without seizing upon the consecutive modes of being.

The two axes underlie not only language as a whole, but also any single linguistic form. On this level, however, the neogrammarian attitude, which eliminates the axis of simultaneity, has survived in the Saussurian doctrine which postulates “the linear character of the signifier (*signifiant*).” The attempt to warrant this “fundamental principle” by a reference to the impossibility of implementing two phonemes at the same time is nothing but a vicious circle, since the phoneme is defined by the impossibility of implementing two such units simultaneously. This arbitrary thesis prevented both its author and the phonemicists following in his tracks from solving the pivotal

problem which he himself astutely foresaw: the task of “determining the distinctive features (*éléments différentiels*) of the phonemes.” The group relations were still examined only in terms of sequences and not at all in terms of bundles. Long ago the neurologists distinguished two kinds of complexes (*Simultankomplexe* und *Sukzessivkomplexe* in K. Kleist’s terminology) which underlie our speech-ability, which are differently located in the brain and which may be respectively compared with the chords and sequences in music. Meanwhile the science of language continued to treat the phoneme as the most minute (further indivisible) linguistic unit.

However, as the phonemes of a given language form a system of sequences, so the system of phonemes, in turn, is formed by their constituents, i.e., by distinctive features. And the breaking up of the phonemes into distinctive features follows precisely the same tested devices as the division of the morphemes into phonemes. “By studying the possible *commutations*” we obtain, e.g., a French phonemic “paradigm” */bu/* ‘boue’ : */mu/* ‘mou’ : */pu/* ‘pou’ : */vu/* ‘vous’ : */du/* ‘doux’ : */gu/* ‘goût’, and thus we find out that the phoneme *b* in */bu/* can be decomposed into five commutable elements: *b/m*, *b/p*, *b/v*, *b/d*, *b/g*. In examining the same phoneme in other environments we confirm this scheme. Cf. */bo/* ‘beau’ : */mo/* ‘mot’ : */vo/* ‘veau’ : */po/* ‘peau’ : */do/* ‘dos’ : */go/* ‘gau’, etc. Having ascertained a set of “micro-features,” as we could call them after the model of Twaddell’s “micro-phonemes,” we may broach the question of the “macro-features,” and we easily detect a set of ratios, as for instance—*/bu/* : */mu/* = */du/* : */nu/* ‘nous’; */bu/* : */pu/* = */du/* : */tu/* ‘tout’ = */gu/* : */ku/* ‘cou’ = */vu/* : */fu/* ‘fou’ = */zuzu/* ‘zouzou’ : */su/* ‘sou’ = */zu/* ‘joue’ : */ʃu/* ‘chou’; */bu/* : */du/* = */pu/* : */tu/* = */vu/* : */zu/* = */fu/* : */su/* = */mu/* : */nu/*; and finally */bu/* : */gu/* = */pu/* : */ku/* = */vu/* : */zu/* = */fu/* : */ʃu/* — cf. */bõ/* ‘bon’ : */gõ/* ‘gond’ = */võ/* ‘vont’ : */jõ/* ‘jonc’ = */mõ/* ‘mon’ : */nõ/* ‘gnon’. In this way the distinctive features constituting the French consonantal phonemes come to light: nasality vs. the absence of nasal resonance; tenseness vs. laxity; continuousness vs. brokenness; gravity (due to a vaster and more unified resonator with a contracted orifice) vs. acuteness; higher vs. lower saturation (respectively due to a back

and a front articulation).² *Entia non sunt multiplicanda*: higher saturation (correspondingly, back articulation) presents one single indivisible feature in the pattern of French consonants, and it is phonemically irrelevant, toward which part of the palate this back articulation is produced; the variants are due to the environment: the articulation is antepalatal when combined with nasality, palato-alveolar in combination with continuousness, and postpalatal elsewhere. Thus the question of contextual variance is quite as pertinent for bundles as for sequences.

Only when brought up to the level of distinctive features, does the linguistic analysis enable us to verify Saussure's cardinal statement on phonemic units as first and foremost "*entités oppositives*." The phoneme by itself is not a term of opposition. E.g., the phoneme *b* does not call unequivocally, irreversibly, and necessarily for a definite opposite, while any distinctive feature does. There is no possible distinctive tenseness without laxity, no gravity without acuteness, no higher without lower saturation, or, as Saussure pointed out, no presence of nasal resonance without its absence, and *vice versa*. The dichotomous principle applies here in full force.

Let us exemplify the results of breaking up phonemes into bundles of distinctive features. The inventory of Standard Serbocroatian totals 29 qualitatively distinct phonemes, and if we add the phonemes distinguished by prosodic features the amount of phonemes swells to 47. The whole pattern is based on eight dichotomous properties; among them six *inherent* (or *qualitative*) features concerning the axis of simultaneity only (vocality, nasality,

saturation, gravity, continuousness, and voicing), and two *prosodic* features involving also the axis of successiveness (length, and high-tone). We mark the presence of a feature by a plus sign; its opposite, the distinctive absence of the given feature, is indicated by a minus. A complex combining both opposite terms is represented by the $\pm$ sign. To avoid longer comments, the current Croatian spelling form is used for denoting the Serbocroatian phonemes.

Prosodic features divide any vowel, as well as *r*, into four distinctive varieties: high-toned long, high-toned short, low-toned long, and low-toned short.³

Thus there must be added to the analytic records of *r* and of any vowel the following four commutable characteristics:

	1	2	3	4
High-tone	-	-	+	+
Length	-	+	-	+

This analytic phonemic transcription, when applied to connected speech-utterances, exhibits the distribution of the different distinctive features throughout the sequences.

On the level of phonemes there are 1081 "significant differentiations" in Serbocroatian, according to W. F. Twaddell's formula: "If *x* is the maximum number of significant phonological differentiations within a given articulatory range in a language, then $2x = n(n-1)$, where *n* is the maximum number of phonemes in that range, and (*n* - 1) is the number of consecutive phonological relations within that range."⁴

In dissociating the phoneme into distinc-

	t	d	c	s	z	p	b	f	v	ć	đ	č	g	š	ž	k	g	x	n	m	ñ	r	l	l'	i	u	e	o	a
Vocality																													
Nasality																													
Saturation	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Gravity	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+
Continuousness	-	-	±	+	+	-	-	+	+	-	-	±	±	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+
Voicing	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-

²Cf. R. Jakobson and J. Lotz, "Notes on the French Phonemic Pattern," *Word* 5 (1949).

³Without an adjacent vowel, *r* is syllabic. Without an adjacent consonant, *r* is non-syllabic. When *r* adjoins both a vowel and a consonant, the phonemic opposition long/short is in force and is implemented as syllabic/non-syllabic. The non-syllabic *r* is prosodically undifferentiated.

⁴*On Defining the Phoneme* (Baltimore, 1935), p. 53.

tive features we isolate the ultimate linguistic constituents charged with semiotic value. When determining their specific essence, do we slip from the linguistic level into physical or physiological criteria and sin thereby against the epigraph of this paper? One must recognize that language obviously pertains to the domain of culture and that even the minutest element charged with semiotic value is a manifestation of culture, for instance, the "functional role" played by the opposition of voiced and unvoiced consonants in Serbocroatian. But as far as the phonic essence of this opposition is concerned, one would be tempted to assign it not to culture but to nature as a purely material phenomenon.

Certainly the existence of sound production with or without glottal vibrations is conditioned by the structure of our vocal apparatus, and the subsequent difference of the acoustic effect is a physical phenomenon. The phoneticians have ascertained that the emission of consonants presents an infinity of degrees and shades with regard to the participation of voice: the glottis can be closed to a greater or lesser degree; the vibrations of the vocal cords can be of different amplitude; and the phase at which they begin or cease can vary. Thus, the glottis is capable of producing diverse nuances in the matter of consonantal voicing, but only the opposition "presence vs. absence of voicing" is utilized to differentiate word meanings.

Since the sound matter of language is a matter organized and formed to serve as a semiotic instrument, not only the significative function of the distinctive features but even their phonic essence is a cultural artifact. Of course, it must reckon with natural limitations; it is, to use Saussure's apt expression, "*en quelque sorte imposé par la nature*." But at the same time our organizing activity superimposes its own rules. Phonemic entities draw on the gross sound matter but readjust this extrinsic stuff, dissecting and classifying it along their own lines. Above all, the procedure is one of selection. Among a multitude of acoustico-motor possibilities, there is a restricted number upon which language chooses to set a value.

Where nature presents nothing but an indefinite number of contingent varieties, the intervention of culture extracts pairs of oppo-

site terms. The gross sound matter knows no oppositions. It is human thought, conscious or unconscious, which draws from this sound matter the binary oppositions for their phonemic use. The term *opposition* (or correspondingly, *contrast*) is currently used in modern linguistic literature, but it is still opportune to recall the vital implications of this concept as, for instance, H. J. Pos neatly formulated them: "L'opposition n'est pas un fait isolé: c'est un principe de structure. Elle réunit toujours deux choses distinctes, mais qui sont liées de telle façon, que la pensée ne puisse poser l'une sans poser l'autre. L'unité des opposés est toujours formée par un concept, qui, implicitement, contient les opposés en lui et se divise en opposition explicite quand il est appliqué à la réalité concrète. . . . L'opposition dans les faits linguistiques n'est pas un schéma que la science introduit pour maîtriser les faits, et qui resterait extérieur à ceux-ci. Son importance dépasse l'ordre épistémologique: quand la pensée linguistique range les faits d'après les principes d'opposition et de système, elle rencontre une pensée qui crée ces faits mêmes."⁵

As music imposes upon sound matter a graduated scale, similarly language imposes upon it the dichotomous scale which is simply a corollary of the purely differential role played by phonemic entities. The consonantal opposition of strong and weak is achieved by varying degrees of air output (*débit* in Rousselot's terminology). Speech sounds present an ample progression in output: the latter is higher in aspirated than in nonaspirated consonants, in fortes than in lenes, in voiced than in unvoiced, in stops than in corresponding spirants. In different positions the relation strong/weak can be implemented by different variants: for instance, in a strong (stressed, initial, etc.) position, by an aspirated fortis stop vs. a lenis stop (voiced or unvoiced), and in a weak position by matching two nonaspirated stops—a fortis and a lenis one, or two spirants—an unvoiced and a voiced one. In Danish this opposition strong/weak is implemented, for example, by *t* vs. *d* in strong position, and by *d* vs. *ð* in weak position, so that the weak phoneme in the strong position *materially* coincides with

⁵"La notion d'opposition en linguistique," *Onzième Congrès International de Psychologie* (Paris, 1938), p. 245.

the strong phoneme in the weak position. The relation strong/weak in any position is perfectly measurable both physically and physiologically, as in general every phonemic opposition presents in all its manifestations, a common denominator both on the acoustic and the articulatory level. But if one should measure the sound matter without reference to the rule of dichotomy imposed upon it by language, the conclusion would be that there are "overlapping" phonemes, in the same way as a physicist with his acoustic instruments, according to H. Frei's felicitous comparison, fails to explain why, in a given piece of music, F-flat and E represent two different values.⁶ The dichotomy of distinctive features is, in essence, a logical operation, one of the primary logical operations of a child and—if we pass from ontogeny to phylogeny—of mankind. The question arises as to whether it is justifiable to admit a kind of logical operation which escapes the attention of the speech-community and which unfolds outside our consciousness. But it is sufficient to recall Ribot's focal thesis: "the reasoning, whether conscious, subconscious or unconscious, remains identical, save in differing degrees of clarity of representation."⁷

And now to return to our starting point: since phonemes are linguistic elements, they are subject to a strictly linguistic analysis which must specify all the underlying oppositions and their interrelations. The specific property of any opposition, the phonic nucleus utilized semiotically, must be exactly identified. Otherwise even a rudimentary listing of the phonemes of a given language becomes a scientifically insoluble problem. What enables us to treat two units appearing each in a different environment ("microphonemes" in Twaddell's conception), e.g., the Spanish initial stop *d* and the intervocalic spirant *ð*, as pertaining to the same phoneme (Twaddell's "macrophoneme")? If we did not submit the phonemes to the analysis sketched above,

what would authorize us to consider these units equivalent? Their identification is then, patently or latently, based on the precarious criterion of outer (physical or physiological) "resemblance," or on an even more wavering criterion of subjective feeling for "sameness."

But these difficulties, particularly striking on the "macrophoneme" level, begin already on the "microphoneme" level, when the "method of commutation" is used. Without the advocated internal analysis of the phoneme one can scarcely know which of the three contiguous variations is a phonemic distinction in such a set of Russian words as [*v'era*] 'belief' (Nom.)—[*v'ér'æ*] 'believing', [*v'eru*] 'belief' (Acc.)—[*v'ér'y*] 'I believe', [*v'eru*] 'belief' (Gen.)—[*v'ér'i*] 'belief' (Dat.).⁸ A reference to the *r* in final position would help somewhat ([*žár*] 'heat'—[*žár'*] 'roast!') but the matching of the intervocalic trill *r* and the more flap-like final *r* returns us to the embarrassing "macrophoneme" problem, and in such a set as [*sláva*] 'glory' (Nom.)—[*sláv'æ*] 'glorifying', [*sláv'u*] 'glory' (Gen.)—[*sláv'i*] 'glory' (Dat.), no reference to other positions is of assistance since the opposition *v/v'* is always prevocalic.

Only in resolving the phonemes into their constituents and in identifying the ultimate entities obtained does phonemics arrive at its basic concept (which insures the consistent use of linguistic criteria sought by Hjelmslev) and thereby definitely breaks with the extrinsic picture of speech vividly summarized by L. Bloomfield: a continuum which can be viewed as consisting of any desired, and, through still finer analysis, an infinitely increasable number of successive parts.⁹

Linguistic analysis, with its concept of ultimate phonemic entities, signally converges with modern physics, which has revealed the granular structure of matter as composed of elementary particles.

⁶The phonetic transcription used is approximate (and without special signs for reduced vowels). The same series as transcribed phonemically: [*v'era*]—[*v'ér'u*], [*v'ér'u*]—[*v'ér'u*], [*v'ér'i*]—[*v'ér'i*].

⁸*Language* (New York, 1933), p. 76.

⁹"Monosyllabisme et polysyllabisme dans les emprunts linguistiques," *Bulletin de la Maison Franco-Japonaise* 8, No. 1 (Tokyo, 1936), p. 146.

⁷See also E. Sapir, "The Unconscious Patterning of Behavior in Society," *The Unconscious* (New York, 1928).

Written in Hunter, N.Y., Summer 1949, and published in *Travaux du Cercle Linguistique de Copenhague* 5 (1949), pp. 205–213, dedicated to Louis Hjelmslev.