Šefčík, Ondřej

Values, features, fine metrics and oppositions

Sborník prací Filozofické fakulty brněnské univerzity. A, Řada jazykovědná. 2008, vol. 57, iss. A56, pp. [5]-14

ISBN 978-80-210-4617-7
ISSN 0231-7567

SBORNÍK PRACÍ FILOZOFICKÉ FAKULTY BRNĚNSKÉ UNIVERZITY
STUDIA MINORA FACULTATIS PHILOSOPHICAE UNIVERSITATIS BRUNENSIS
A 56, 2008 — LINGUISTICA BRUNENSIA

ONDŘEJ ŠEFČÍK

VALUES, FEATURES, FINE METRICS AND OPPOSITIONS

(This paper was written with the help of Grantová agentura ČR, # 405/06/P179)

In our former paper (Šefčík 2007: 23–24) we defined the term **fine metrics** (in Czech **jemná metrika**), which was applied both on a set of phones and on a set of phonemes. A brief definition of fine metrics will be given below.

In the present paper we would like to add to our former reading some addenda concerning terms like *value, feature* and to show some practical solutions of applying fine metrics on phonemic material. The final remarks will consider some notes on the relationship between fine metrics and the opposition-theory as defined by Trubetzkoy, Cantineau and Marcus.

The very first step in the current paper is a definition of the term *value*, followed by an example how such values could be used as a set of components, which could be easily metrized, and how such a metrized phonemic system could be easily explained by such metrics, according to the above-mentioned theory of oppositions.

Note 1: All examples will be taken from Vedic Sanskrit (= Old Indo-Aryan).

Values of phones and phonemes

We should keep in mind that speech material is linear, indiscrete and continual mass of speech-events hence the segmentation itself demands a certain degree of objectivization of the material.

Identification of phones as discrete segments of given strings demands on a phonologist a postulation of criteria for classification that are used for distinguishing given phones in strings from each other. It is only when all phones are identified, a phoneme analysis can begin.

The identification of phones or phonemes is always based on the set of values of a given phonemic segment; without such a description of values used, identification is not possible at all.

On the following lines of the present paper we will deal only with phonemes and values of phonemes; however, it should be kept in mind that values of phonemes are derived from values of phones.

On the other hand, an analysis of phonemes and their systems have much higher priority for any phonologist, hence our reading will be focused on phonemes. In fact, all of the following observations could be applied on phonemes as well as on phones.

Values and their classification

Values of given phonemes are here considered as ordered sets attached to concrete phonemes of a given phonemic system.

Any phoneme (according to common tradition) is written between slash brackets, for example */x/*. Any value of a given phoneme is here symbolized $v_i/x/$ in general or in the case of a concrete value simply as */value/*. The sum of values of any phoneme is symbolized as $V/x/$ (i.e. $V/x/ = v_1/x/\cup v_2/x/...\cup v_i/x/$), the set of values of all phonemes of a given language is symbolized simply as V, which is the set union of particular values of concrete phonemes, hence $V = V/x/_1 \cup V/x/_2 ...\cup V/x/_i$).

Any value could be classified with the help of three criteria, every one expressing some properties of the given value in the phonemic system. Each criterion consists of a pair of possible incompatible properties.

The first criterion is of **homogeneity**. Two values are **homogeneous** if they are members of the same subset $V_{i,}$ otherwise such values are **heterogeneous** (Marcus 1967: 46–47, Marcus 1969: 50, Kortland 1972: 57).

Example 1: Values */voiced/* and */unvoiced/* are members of the same subset V_1 in Vedic Sanskrit, hence they are homogeneous.

Example 2: Values */voiced/* and */lengthened/* are not members of the same subset V_2 in Vedic Sanskrit, hence they are heterogeneous.

The second criterion is of **compatibility**. Two values are **compatible** if they are attached to one phoneme */x/* (members of the same set of phoneme values $V/x/$). If such values are not members of the same set of phoneme values (they are not attached to the one phoneme), they are **incompatible** values. It should be mentioned that all compatible values are heterogeneous, but not otherwise, so some heterogeneous values are not compatible (Marcus 1967: 47–48, Marcus 1969: 50–51, Kortland 1972: 57).

Example 3: Values */voiced/* and */aspirated/* are in Vedic Sanskrit compatible.

Example 4: Values */voiced/* and */unvoiced/* are in Vedic Sanskrit incompatible.

Example 5: Values */voiced/* and */aspirated/* are in Vedic Sanskrit compatible, however heterogeneous.

The third criterion is of **contrastivity.** Any values v_i and v_j are **contrastive** if there are two phonemes such that $V/x/ - V/y/ = v_i$ and $V/y/ - V/x/ = v_j$. In other words, if replacement of one value of a given phoneme by another value results in another phoneme, the values are contrastive. Otherwise both values are **incontrastive**. All contrastive values are homogeneous and therefore incompatible, but not otherwise (Marcus 1967: 48–49, Marcus 1969: 51–53, Kortland 1972: 58).

Example 6: Vedic phonemes */t/* and */d/* are contrastive, as the complement of *V/d/ – V/t/ = /voiced/* and *V/t/ – V/d/ = /unvoiced/*.

Note 2: Revzin (1966: 18–20) does not distinguish contrastivity from homogeneity (such value is by him called homogeneous). However, it was Revzin, who inspired Marcus to his classification.

Values and their expression using the code

It is clear from what was written above that values could be considered as coordinates in the phonemic space and the set of values of any phoneme could be expressed as a vector with values as coordinates.

Such a set of phonemic values is then expressed by a code, which is constructed in the following way.

The axes of a considered phonemic space are designed by compatible values. All compatible values are then ordered in a string where every position is attached to concrete compatible values.

A pair of homogeneous incompatible values is expressed by numbers present in a given position. From possible homogeneous values, we prefer those that are contrastive, hence we can use only binary marking. Then zero (*0*) is assigned for those values, which are unmarked in the sense of Trubetzkoy and other scholars. A marked value will be symbolized by the number one (*1*).

Every pair of contrastive, homogeneous and incompatible values we term a (distinctive) **feature**. For example, the pair of contrastive homogeneous incompatible values *V/d/ – V/t/ = /voiced/* and *V/t/ – V/d/ = /unvoiced/* form the feature */±voice/*.

It is necessary to keep in mind that in every language the features are determined by values, which are in that language actually used. Features are not universally established, as both Jakobsonian or generative phonologists think, but features should be independently found and described properly for any analyzed language.

Note 3: For history and today practice of the use of the term *feature* see Jakobson – Fant – Halle 1952: 8–15, Jakobson – Halle 1956: 3–6, 33–36, Chomsky – Halle 1968: 64–69, 164–170, Baltaxe 1978: 10–19, 47–68, Akamatsu 1988: 77–110, Hall 2001: 1–40, de Lacy 2006: 57–59, 247–249, Steriade 2001: 139–157 (in de Lacy 2007 (ed.)).

Note 4: Marcus (1967: 51, 1969: 54–55) uses the term “feature” (i.t. “trait”, “rys”) for non-homogeneous value! His term, equivalent to the common view of the feature is “value pertinent” (1967: 55–58, 1969: 58–61).

On binarity of feature

In most cases of phonemic analysis a set of values *V/x/* is set up by a code formed by only contrastive homogeneous values, i.e. of binary quality marked by *0* and *1*.

Example 7: Vedic vowels could be described by using a feature /±*length*/, hence every vowel phoneme has value of either /-*length*/ or /+*length*/.

Example 8: Vedic stops could be described by using features /±*voice*/ and /±*aspiration*/. Hence every stop phoneme has values /-*voice*/ or /+*voice*/, /-*aspiration*/ or /+*aspiration*/, respectively.

Unfortunately, a real phonemic system is not always in the accord with the ideal. Some of values are incompatible and homogeneous, but not contrastive.

Example 9: Vedic Sanskrit has three basic local series of stops: velar, dental and labial (in addition to them there are two related series yet – see Šefčík 2002). In the terminology of classical phonology, such basic series are in an equipollent opposition (see below) and the location of any series cannot be classified as contrastive, just for that reason that the fact of the given series not being velar does not necessary mean that series is dental (it could be labial) and vice versa.

There are two possible solutions of such a problem. The first solution is to incorporate three-valued features and by that means to free the given position in a string of compatible values from the property of contrastivity but to preserve the property of homogeneity. Then we can mark the dental value of "localization feature" as *0*, the velar value as *1* and the labial value as *2*. The advantage of such a solution is in keeping for localization one and only position in chains but on the other hand it defects the elegancy of a phonemic description (there are not only contrastive features, but there are incontrastive "features", too). Another disadvantage is that such a solution leads towards a false understanding of dental series as unmarked against two marked series (of velar and labial), which does not correspond to the reality of an equipollent opposition.

Another solution is traditionally binary: an equipollent opposition is carefully disjoint on the set of privative (and binary) oppositions by the introduction of two features (based on some subphonematic property) which are applied on the subsystem – for example Jakobsonian /*compact-diffuse*/ and /*grave-acute*/. Then dental series is marked *00*, labial *01* and velar *10*. A great advantage is that all of the used features are in addition contrastive. But the disadvantage of false understanding of marking remains.

Our solution is closer to the second one in disjoint of local incontrastive "feature" on more contrastive features, but we differ in introducing of three features, hence none of the considered series could be misunderstood as unmarked, all of them are marked by one marker. We will call such introduced features **bounded features** and we will mark them in the code by underlining. The whole process of analysis is that ordered in a series of steps.

First, we have one incontrastive location feature with three values of dental (*0*), velar (*1*) and labial (*2*) location as above (a capital letter expresses any stop of a given location, other values are omitted and marked only by *x*):

V/P/	=	/labial/	=	0	in code /xx0xx/
V/T/	=	/velar/	=	1	in code /xx1xx/
V/K/	=	/dental/	=	2	in code /xx2xx/

Second, we construe given homogeneous values as a triple of orthonormal vectors. Accordingly, any labial has the value *1*on the labial axe and *0* on other local axes. Similarly, velar has the value on the velar axe *1* and *0* on other local axes and dental has the value *1* on the dental axe and *0* on other local axes. See the following codes:

V/P/	=	/labial/	=	001	in code /xx001xx/
V/T/	=	/dental/	=	010	in code /xx010xx/
V/K/	=	/velar/	=	100	in code /xx100xx/

Notice bounded value /xx000xx/. It could be taken for granted as unmarked; it is not occupied by any real phoneme (just for that reason that all Vedic stops have location mark).

Practical remarks on the work with codes

It is clear from the above written examples that to work with differences between two or more sets of phonemic values given by an *n*-position in the chain could be very difficult simply for practical reasons – codes can be long and a reader could not easily understood in which position in the code two or more phonemes differ or what feature or value a given position expresses.

Such practical problems are hardened by the fact that we usually compare two very similar codes belonging to two very similar phonemes, often different simply on one and only one position. All other features/values are then for reasons of description irrelevant and they can be omitted just as we did above. Our solution was that such redundant features were replaced by a common symbol *x*, but it was not an elegant solution.

A better solution could be to look at the very structure of codes of the compared phonemes. The part of a code, which is same for all sets of values, is the intersection of values of all considered phonemes (or in other words – which is common for all considered phonemes), is *base of opposition* (see below). The part of code, which is not common for all considered phonemes, is then *comple-*

ment of opposition. So we can remove a whole part of code, which is the base of opposition, out of the code and mark such segment. A remaining feature (or features), which forms a complement, can be then compared independently.

Example 10: Let us have two sets of values, f. e. *V/t/* and *V/d/*. When these two sets of values are compared, they differ in only one position, in position of the feature /±*voice*/, all other values are same. Those values are then the base of opposition and we can mark this base as a /*T*/. Let the value /-*voice*/ be marked as *0*, the value /+*voice*/ as *1*. Than the set of values *V/t/* can be rewritten as $^{\text{voice}}0/T/$ and set of values *V/d/* as $^{\text{voice}}1/T/$.
If we deal with more than one feature, for instance, in addition to /±*voice*/ with the feature /±*aspiration*/, than we can work out accordingly the above example in the following way: $V/t/ = {}^{\text{voice}}0^{\text{aspiration}}0/T/$, $V/d/ = {}^{\text{voice}}1^{\text{aspiration}}0/T/$, $V/t^h/ = {}^{\text{voice}}0^{\text{aspiration}}1/T/$, $V/d^h/ = {}^{\text{voice}}1^{\text{aspiration}}1/T/$.
It could be useful to simplify the record by omitting names of features, if they are clear from the context. Then the record of codes could be simply $V/t/ = 00/T/$, $V/d/ = 10/T/$, $V/t^h/ = 01/T/$, $V/d^h/ = 11/T/$.

Values and distance

The distance between phonemes is given by the distance between given *n*-tuples of sets of values of given phonemes. A distance between any such *n*-tuples can be metrized using fine metrics.

Any metrics is defined as a tuple (A, ρ), where A is a set of components and ρ is a distance on A, or a function $(A \times A)$, such that the metric space has the axioms of 1) identity, i.e. $\rho(x, y) = 0,$ if $x = y;$ 2) symmetry, i.e. if $\rho(x, y) = \rho(y, x)$ and 3) triangle inequality, i.e. $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$ (see Marcus 1967: 34–35, Marcus 1969: 42, Brainerd 1971: 90, Šefčík 2007: 20).

Fine metrics must comply with the above written axioms. The set of components is a set of values of phonemes of a given language. A distance is given by the complement between such sets of values.

As was already shown above, such a set of values of given phonemes could be expressed as a code. The distance between codes can be metrized as the Hamming distance, given by the number of differences between codes.

Comparing two (or more) codes, we mark the match between compared codes on a given position by *0* (= zero difference), if there is no match in the same position, we mark *1*. After that, we count the number of *1*s. The sum expresses the distance between sets of values of phonemes, i. e. between the phonemes themselves.

Example 11: The distance between sets of values equal to *1*:

V/t/	00/T/	
V/d/	10/T/	
V/t/ – V/d/	100	Σ =1

Because the sets of values of phonemes /*t*/ and /*d*/ differ only in one feature (i.e. /±*voice*/), the distance between both phonemes is equal to *1*.

Example 12: The distance between sets of values equal to *2*:

V/t/	00/T/	
V/d^h*/*	11/T/	
V/t/ – V/d^h*/*	110	Σ = 2

Because the sets of values of both phonemes differ in two features (i.e. /±*voice*/ and /±*aspiration*/), the distance between both phonemes is equal to *2*.

Values and oppositions

The classical theory of oppositions, as developed by Trubetzkoy (1939) and reworked by Cantineau (1952, 1955) and Marcus (1967, 1969), can be, for its universality, applied on values, features and metrics above them, too.

Note 5: In the comment below we differ from our predecessors with a more specified definition of disjunctive opposition.

Two basic ideas of the theory of opposition are terms **base (of opposition)** and **complement (of opposition)** (see Marcus 1967: 9–10, Marcus 1969: 22, Brainerd 1971: 20–22). Here we give only a short overview of basic terms of theory of opposition with a pointed-out relation to the terms of value, feature and fine metrics.

For us, the base of opposition equals the same values (of the features of the features of given phonemes) present in both (all) sets of values of compared phonemes. Or in other words, the core is the intersection of given sets of values, the common part of both (all) phonemes (i.e. *V/x/* ∩ *V/y/*).

The complement of opposition are those values (of the features of given phonemes) not present in both (all) sets of values of compared phonemes. In other words: the complement is that part of values of given phoneme which is not common to another phoneme or other phonemes (i.e. *V/x/* - *V/y/* or *V/y/* - *V/x/*).

Let us classify oppositions and try applying fine metrics on them.

Note 6: In lines below the theory of opposition will be presented in the form closest to Marcus (1967, 1969) and Brainerd (1971), older citations have only a historical value.

The first criterion of classification of oppositions is that which deals with classification of relations between oppositions.

If the complement of the opposition between sets of values of phonemes *V/v/* and *V/w/* is equal to the complement of the opposition between sets of values of phonemes *V/x/* and *V/y/* (i.e. *V/v/ – V/w/ = V/x/ – V/y/)*, both oppositions are **proportional**. If the opposition between sets is not proportional with another, such an opposition is **isolated** (see Trubetzkoy 1939: 63–66, Marcus 1967: 12–13, Marcus 1969: 24–25)**.** It is obvious that if the distance between *V/v/ – V/w/* is same as the distance between *V/x/ – V/y/*, so the distance between *V/v/* and *V/w/* is equal to the distance between *V/v/* and *V/w/*.

Example 13: Vedic phonemes /*t*/ and /*d*/ are in proportional opposition to /*p*/ and /*b*/, because *V/t/ – V/d/ = /+voice/ = V/p/ – V/b/*. The distance between /*t*/ and /*d*/ is equal to the distance between /*p*/ and /*b*/.
Example 14: Vedic phoneme /*a*/ is in isolated opposition to the phoneme /*s*/.

If the opposition between *V/v/ – V/w/* has the same base as the opposition between *V/x/ – V/y/*, both are in **homogeneous** opposition. If there is no same base they are in **singular** opposition (see Trubetzkoy 1939: 60–63, Marcus 1967: 16–17, Marcus 1969: 28–29).

Example 15: Vedic phonemes /*t*/ and /*d*/ are homogeneous to /t^h/ and /d^h/.

Oppositions can be classified according to the criterion of a relation between its members, too.

The very first possibility is such that if a set of values *V/x/* is equal to the set of values *V/y/* and vice versa. The sum of differences between codes is then a void set and the distance between phonemes is null, hence both sets belong to the one phoneme. Or in other words, the base of opposition is equal to both phonemes and the complement is void. This kind of opposition is **null opposition** (see Marcus 1967: 4–7, Marcus 1969: 17–22, Brainerd 1971: 20–22)**.**

Although such a type of opposition is for a phonologist of relatively small importance, it must be kept in mind that such oppositions always exist.

But for phonologists those types of oppositions are of much importance that are not null.

If two sets of values *V/x/* and *V/y/* have the common base of opposition except values of one and only one feature and the complement is null for one of the phonemes (i.e. null value of a given feature) and *1* for the second of the phonemes (i.e. non-void value of a given feature), then such a kind of opposition is known as **privative** (see Trubetzkoy 1939: 67, Marcus 1967: 4–7, Marcus 1969: 17–22, Brainerd 1971: 20–22). The Hamming distance between *V/x/* and *V/y/* is equal to *1*.

Example 16: Phoneme /*t*/ differs from phoneme /*d*/ only in the feature /±*voice*/, all other features are same.

If two sets of values have the common base and the complement is non-void for both phonemes, then the opposition is **equipollent** (see Trubetzkoy 1939: 67, Marcus 1967: 7–9, Marcus 1969: 20–21, Brainerd 1971: 20–22). The Hamming distance between *V/x/* and *V/y/* is equal to *2*.

Example 17: If *V/T/ = <u>001</u>/STOP/, V/K/ = <u>100</u>/STOP/* and *V/p/* is *<u>010</u>/STOP/*, then distance between any pair of values is equal to *2*.

Finally, if any set of values *V/x/* contains some features and a set of values *V/y/* contains at least one feature which is not present in *V*/*x*/, then such sets of values of phonemes are in **disjunctive opposition** (see Marcus 1967: 7–9, Marcus 1969: 20–21, Brainerd 1971: 20–22). The distance between such sets cannot be metrized using Hamming distance, because Hamming distance can be applied only on sets with same number of digits.

Final remarks

In the present paper we tried to sketch how the term of fine metrics could be useful in a phonological research, especially if it is applied together with the terms value and feature, but it could be successfully used in a theory of opposition, too.

Fine metrics is another tool of a more perfect analysis of the phonological system, but it is not (it should not be) one and only one method used.

LITERATURE

AKAMATSU, T. (1988): *The theory of neutralization and the archiphoneme in functional phonology*. Armsterdam: John Benjamins

BALTLAXE, C. A. M. (1978): *Foundations of Distinctive Feature Theory*. Baltimore: University Park Press

BRAINERD, B. (1971): *Introduction to the Mathematics of Language Study*. New York: American Elsevier

CANTINEAU, J. (1952): Les oppositions significatives. *Cahiers Ferdinand de Saussure* 10, 11–40

CANTINEAU, J. (1955): Le classement logique des oppositions. *Word* 11/1, 1–9

CHOMSKY, N. – HALLE, M. (1968): *The Sound Pattern of English*. New York: Harper & Collins

HALL, T. A. (ed.) (2001): *Distinctive Feature Theory*. Berlin: Mouton de Gruyer

HUBEY, H. M. (1999b): *Mathematical and Computional Linguistics*. Muenchen: Lincom Europa

KORTLAND, F. H. H. (1972): *Modelling the Phoneme – New Trends in East European Phonemic Theory*. The Hague: Mouton

DE LACY, P. (2006): *Markedness. Reduction and Preservation in Phonology*. Cambridge: Cambridge University Press

de Lacy, P. (ed.) (2007): *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press
MacDonell, A. A. (1910): *Vedic Grammar*. Strassburg: Karl J. Trübner
MacDonell, A. A. (1916): *A Vedic Grammar for Students*. Oxford: Clarendon Press
Marcus, S. (1967): *Introduction mathématique ŕ la linguistique structurale*. Paris: Dunod
Marcus, S. (1969): *Algebraické modely v lingvistice*. Praha: Academia
Revzin I. I. (1966): *Models of Language*. London: Methuen & Co.
Steriade, D. (2007): Contrast. In: de Lacy (ed.) 2007. 139–157
Šefčík, O. (2002): *Trubeckého* Schwesterreihen *a jejich syntagmatické alternace*. SPFFBU A 50. 99–111
Šefčík, O. (2007): *K užití metriky ve fonologii*. SPFFBU A 55. 19–26
Trubetzkoy, N. S. (1939): *Grundzüge der Phonologie.* Prague: TCLP

HODNOTY, RYSY, JEMNÁ METRIKA A OPOZICE

V tomto příspěvku se věnujeme některým otázkám, spojeným s pojmem jemné metriky, tj. takové metriky, jakou použijeme na vyjádření vzdálenosti mezi množinami hodnot různých fonémů.

V článku vycházíme z pojmu hodnota, kterou klasifikujeme podle kritérií homogennosti/heterogennosti, slučitelnosti/neslučitelnosti a kontrastnosti/nekontrastnosti, ukazujeme, jak od tohoto pojmu lze odvodit pojem rysu (jako sjednocení dvou kontrastních, homogenních a neslučitelných hodnot) a jak rysy mohou být vyloženy pomocí jemné metriky.

Na závěr jsou stručně načrtnuty vztahy mezi jemnou metrikou a klasickou teorií opozic, jak ji vyložili Trubeckoj, Cantineau a Marcus.

Příklady jsou převzaty z védského sanskrtu.

Ondřej Šefčík
Ústav jazykovědy a baltistiky
Filosofická fakulta
Masarykova universita
Brno 602 00, A. Nováka 1
e-mail: sefcik@phil.muni.cz