

**ASPECTS OF CHINESE PHONOTACTICS AGAINST
A COMPARATIVE BACKGROUND OF POLISH
DUBIETIES AND PROSPECTS
(*A preliminary draft*)**

JERZY BAŃCZEROWSKI

*Appearances can be deceptive,
in language reality as well*

1. Introductory remarks

The present paper aims to call attention to the possible applicability of a phonotactological theory to the description of a certain fragment of Chinese phonotactics. However, before we proceed to the presentation proper of this theory, let us say in advance a few introductory words as groundwork.

The phonic and graphic universes of ethnic languages, consisting of signs of various length and complexity, are immense. The communicative manageability of these universes is secured among other operations by:

- (i) equalization, and
- (ii) differentiation (distinguishing).

It is by virtue of the former that language signs are treated as being the same, whereas by virtue of the latter as totally or partially distinct.

Phonic and graphic equalization lies at the basis of the reproducibility of lingual signs, and without this operation language communication would be impossible. Phonic and graphic differentiation, in turn, guarantees the availability of a necessary diversity of these signs. The essence and, at the same time, the economy of the phonic and graphic coding of lingual signs consist in the possibility of creating a multitude of sign-

types out of a relatively small number of initial unit-types by the appropriate composition of these latter within their respective phonotaxis and graphotaxis.

Speaking more precisely, the sign-types in question represent the corresponding sign-tokens which, in turn, consist of the atomic unit-tokens, that is, sounds (phonons), and graphons, being ordered sequentially or simultaneously and sequentially. These unit-tokens are represented by unit-types, that is, by phones and graphs, respectively.

The principle of economy of phonic and graphic coding is at work practically in all phonotactic and graphotactic systems. The latter of these are usually called ‘writing systems’, and the principle in question is always in action in phonographic, syllabographic as well as in morphemographic (semiographic) systems. This last system operates in Chinese the orthography of which utilizes characters (ideographs) which represent morphemes or words if they coincide with morphemes.

For theoretical considerations on the phonotactic reality of ethnic languages it seems to be advisable to make a distinction between:

- (i) phonotactology, and
- (ii) phonotactics.

The former conceived of as a subdiscipline of linguistics is a class of linguistic theories and calculi, the subject matter of which happens to be phonotactic reality. The image of this reality projected by phonotactology takes on the shape of the phonotactological domain or, simply, phonotactics (phonotaxis). This latter is comprised of all phonotactic objects and relations, that is, it forms a certain system. Consequently, we can say that phonotactology is about phonotactics or that it treats of phonotactics by virtue of describing and explaining properties of its objects or by generating their representations.

Our subsequent phonotactological inquiry is intended to be restricted to words only, and its goal is, as already mentioned, to propose in broad outline a general phonotactological theory as well as to examine whether it is applicable to describe a certain aspect of Chinese word phonotactics. This theory seems to work relatively well, if applied to words in Polish. But, since the structure of Chinese words diverges in many significant respects from Polish, the author could not resist the temptation to also include Chinese in his investigation. And, besides, if this theory would aspire to a claim of panglottal validity it cannot leave Chinese data unconsidered.

The restriction of our analysis to words, in order to operate with comparable units, causes certain problems, some of them serious. The existence of these lingual units in many languages as, for example, European ones, is less objectionable than the existence of words in Chinese. Some linguists even deny completely the relevance of the concept of word to Chinese, asserting that this language has no morphology. However, without going more deeply into this problem, let us only state that the existence of words in Chinese is for us beyond doubt and in this respect we agree to a considerable extent with Packard (2001: 7ff).

As far as the general phonotactological theory to be outlined here is concerned, it is our conviction that its presentation would be facilitated, if the terms it utilizes, were explained and exemplified with the help of language material taken not only from

Chinese but also from other languages. It is therefore that we decided to avail ourselves of Polish to this end.

Serious difficulties for the exemplification of some of the theoretical concepts with which we shall operate in the course of our study are caused by the unavailability of suitable dictionaries which should satisfy the following requirements:

- (i) they should be sufficiently representative of the vocabulary of a given language,
- (ii) their entries are solely words, that is, they do not contain syntagms composed of two or more words as their entries,
- (iii) they are accessible in an electronic form, and
- (iv) they give their word-entries in phonetic transcription.

The calculations worked out for the purposes of this article by Prof. Piotr Wierzchoń are based on the following dictionaries:

- (i) Doroszewski, W. (ed.). 1958–1969. *Słownik języka polskiego*. (Dictionary of the Polish language). Warszawa: Wiedza Powszechna – PWN. (SJP-D),
- (ii) Grzegorzczkowska, R., Puzynina, J. (eds). 1973. *Indeks a tergo do słownika języka polskiego pod redakcją Witolda Doroszewskiego*. Warszawa: PWN,
- (iii) Wołosz, R. (www.mimuw.edu.pl/polszczyzna/a_terSPAN),
- (iv) *MDBG Chinese-English dictionary*. (CC-CEDICT).

The Polish dictionary satisfies the conditions (i)–(iii). Sporadically, there occur here also syntagm-entries, but they were easily eliminated from the calculations. However, the Chinese dictionary satisfies only conditions (i) and (iii). Its entries given as ideographs (characters) and transliterated in *pinyin* are not only words but also syntagms and the effective exclusion of these latter from the calculations was not possible. Therefore these calculations will operate here with dictionary-entries rather than with word-entries.

Since neither of these dictionaries gives their entries in phonetic transcription, the exemplifications which will be adduced, reflect the graphotactic structure of these entries rather than the phonotactic. Nevertheless, some interesting, although only approximate inferences, can be made. Thus, it remains to be hoped that in the near future the dictionaries satisfying the conditions enumerated above, will be available, which will make it possible to obtain exact data on the phonotactic structure of languages.

Let me still mention that my original intention was to present the theory, proposed here for the attention of those interested, as an axiomatic system. However, since such an undertaking would require making use of a formal apparatus borrowed from symbolic logic, set theory, theory of relations, and mereology, and since the explanation of this apparatus would take more space than is available, I decided to leave an axiomatic formulation of this theory to some future occasion.

2. Some terminological remarks

The conceptual content of our theory will be mirrored in the terms being utilized, some of which will play a special role. To these terms belong word, phone and phonotacteme among others and they will be, in turn, instrumental in introducing some

other terms, fundamental for the theory in question. In order to briefly explain the properties of lingual objects denoted by these three terms let us begin with an utterance, treated here as a unit of ultimate reference *sui generis*. Although subsequently we shall be not directly interested in utterances, they are lingual units from which, through the operation of quasi-segmentation, all other linguistically relevant units are obtainable.

An **utterance** will be understood here as a spatio-temporal physical object, individual and concrete, produced *hic et nunc* by a definite speaker, in a definite time and space. Its existence does not exceed much the time of its production, whereupon disintegrating it vanishes irretrievably into the past. In a certain sense an utterance is a linear object consisting of phonical substance, having its beginning, duration and termination in time, and immediately preceded and succeeded by pauses. In language communication an utterance functions as a sign, simple or composite, which designates and signifies the corresponding extralingual entities. On the one hand, it forms a certain integral, indivisible communicative whole, while on the other hand it is quasi-divisible into linearly quasi-disjunct parts or segments. This quasi-segmentation is performed by native linguators or linguists and results in units of various kinds, among which actual words (vocalons), sounds (phonons), syllables, etc. can be distinguished.

The unit termed here a **vocalon** or an **actual word** is conceived of as a maximal unit of linear, that is, sequential, ordering of an utterance. Putting it differently, the linear structure of an utterance may be imagined as a sequence consisting of vocalons as always linearly continuous and relatively easily distinguishable units within utterances. Obviously, each vocalon as a constituent part of the corresponding utterance is as individual and concrete as the latter. What is more, each vocalon always functions as a sign. Although vocalons are treated as mereological wholes, units of various kinds may nevertheless be distinguished within them. And, this statement leads us straight to the concept of phonaton.

A **phonaton** will be conceived of as any subvocalonic part or segment of various size, provided it is linguistically relevant. Each phonaton is also as individual and concrete as its corresponding vocalon and it is always a linearly continuous unit. Needless to say, every vocalon will be treated as a particular kind of phonaton. Thus, every vocalon is also its own subphonaton. Within the set of all phonatons two kinds may be distinguished: **proper** and **virtual** (vacuous, improper). The former consist of phonic substance, and the latter are asubstantial objects, that is, substantially zero-segments or, simply, moments of silence.

Phonatons may be equal with respect to certain phonetic properties, and distinct with respect to others. Auditory equality or indistinguishability will be conceived of as **homophony**. This relation will be instrumental in defining the concept of phone and vocable.

Every phonaton, if it is not already minimal, may undergo further segmentation. A particular kind of phonaton will be called a **sound** or **actual phone**. However, for technical reasons we shall prefer to use the term **phonon**. This unit is distinguished just as a minimal phonaton which, being the only subphonaton of itself, is not composed of any smaller subphonatons.

The set of phonons (sounds) may be classified into phones. A *phone* can be defined as a set of all those phonons which are homophonous with a given phonon. As can be rightly inferred, any two different phones are disjoint sets of phonons, and each phonon belongs to exactly one phone.

The set of vocabulons (actual words) may be classified into vocables. A *vocable* will be defined as a set of all those vocabulons which are homophonous and homosignificative with a given vocabulon. Consequently, any two different vocables are always disjoint sets of vocabulons. It also becomes clear that the term *word* is ambiguous, since it may mean either a vocabulon or a vocable in our theoretical system.

The vocabulons as characterized above have emerged as certain wholes in the mereological sense. Examining them under a purely physical angle no natural articulatory or acoustic boundaries between the intravocabulonic phonons can be discovered. Nevertheless, native linguators, drawing upon their language knowledge, have a compelling auditory impression that during speaking and listening they, respectively, utter and hear successions of sounds (cf. Jones 1950: 2). Thus, linguators' language knowledge lets them impose an auditory quasi-segmentation upon vocabulons.

In order to formally account for the auditory virtual segmentation of vocabulons we shall associate with them two kinds of phonetic representation of their linear (sequential) structure. In consequence, we shall arrive at objects called phonotactons and phonotactemes, respectively. The former will represent this structure in terms of phonons, and the latter in terms of phones. The *phonotacton* associated with the corresponding vocabulon is a sequence of phonons (sounds) which are constituents of this vocabulon and such that the consecutive members of this sequence reflect the temporal succession of phonons in this vocabulon, and the first member is the initial, and the last member is the terminal pause. Let us still draw attention to the fact that phonons as actual objects cannot repeat themselves in phonotactons. Consequently, each phonotacton is always a sequence without repeated members.

The phonotactons can be converted into phonotactemes by substituting phones for the corresponding phonons. Thus, the substitution of phones for the corresponding phonons in the phonotacton associated with a given vocabulon will result in the *phonotacteme* for this vocabulon. Hence, each phonotacteme will always be a sequence of phones, the first and last member of which are pauses. Of course, phones may repeat themselves in phonotactemes. Usually, the phonotactemes are represented graphically in terms of phonetic transcription. However, for certain technical reasons, we shall utilize capital letters to this end.

3. The concept of tactophoneme

Phonotactemes have been conceived of as sequences of phones representing the corresponding phonotactons, including vocabulons. Certain phones are thus concatenable or tactifiable together in various ways to form phonotactemes. Putting it differently, we can say that such phones are fit for various sequential orderings which result in those

phonotactemes which, in turn, function as representations for the corresponding vocabulons (words) or their constituent phonatons in a given language.

For a description of certain aspects of the concatenability of phones into phonotactemes the concept of tactophoneme seems to be instrumental. By the *tactophoneme* we shall understand a subset of phones the sequentialization of which results in a phonotacteme. However, one and the same tactophoneme may be sequentialized in various ways whereby various corresponding phonotactemes are created. In order to exemplify this operation let us avail ourselves of Polish and Chinese.

Ex 3.1 The Polish tactophoneme (A, K, R) , consisting of three phones, by virtue of all permissible permutational sequentializations results in the following eight phonotactemes: *AKR* (acre), *ARAK* (arrack), *ARKA* (the Ark), *KARA* (punishment), *KARK* (nape of the neck), *KRAK* (a legendary king of Cracow), *KRA* (ice floe), *RAK* (crayfish, cancer).

These eight phonotactemes thus represent the corresponding vocabulons (words) in terms of the sequences of phones. Let us still emphasize that each of these phonotactemes consists exactly of all and only the phones creating the tactophoneme (A, K, R) . Consequently, we can say that (A, K, R) is the tactophoneme for each of the phonotactemes in question.

Ex 3.2 The Chinese tactophoneme (A, A, B, F, N) consisting of five phones, sequentializes only into the following two phonotactemes: *BANFA* (way, means, measure) and *FABAN* (deal with according to law). Also the Chinese six-phone tactophoneme (A, A, B, F, N, W) sequentializes into two phonotactemes: *FANWAN* (rice bowl) and *WANFAN* (supper, dinner).

The Chinese phonotactemes adduced above represent the corresponding compound words (vocabulons) in terms of the sequences of phones. But, at the same time, they represent these words in terms of the sequences of morphs or, more exactly, morphotactemes or, more correctly, phonomorphotactemes.

Thus, Polish avails itself to a great extent of the permutational sequentialization of phones belonging to a tactophoneme to form various simplex word phonotactemes. Chinese, in turn, prefers the sequentialization, or limited permutational sequentialization, of morphs belonging to a tactomorpheme to form composite word morphotactemes (phonomorphotactemes).

What is very conspicuous is the fact that the Chinese tactophonemes produce much less simple word phonotactemes than the Polish tactophonemes. The strong distributional restrictions imposed upon the cooccurrence of phones within morphemes limit the number of permutational sequentializations.

The tactophoneme may be characterized in terms of such properties as:

- (i) phonicity,
- (ii) phonotactemic range, and
- (iii) phonotactemicity.

Particular vocabulonic tactophonemes may be comprised of a differing number of phones, that is, they may be monophonous, biphonous, triphonous, etc. By the phone power of a tactophoneme or, simply, by its *phonicity* we shall understand the number of phones which are its elements.

Ex 3.3 The phonicity of the Polish tactophoneme $\{A, K, R\}$ is thus 3, and the phonicity of the Chinese tactophoneme $\{A, A, B, F, N\}$ is 5. The phonicity of the vocabulonic tactophonemes in Polish ranges from 1 up to 16. Thus, the minimal number of phones of which the tactophonemes in this language are comprised is 1, and the maximal number 16. The determination of phonicity of vocabulonic tactophonemes in Chinese is more difficult, because the identification of maximal vocabulons (words) causes problems.

The set of all phonotactemes which result from the sequentializations of the same tactophoneme will be termed the *phonotactemic range* of this tactophoneme.

Ex 3.4 All the eight phonotactemes created out of the tactophoneme $\{A, K, R\}$ in Polish form its phonotactemic range. Similarly, the two phono-tactemes created out of the tactophoneme $\{A, A, B, F, N\}$ in Chinese form its phonotactemic range.

In addition to the concept of phonotactemic range of a tactophoneme, we shall also operate with the *phonotactemic onus (load)* or, simply, the *phonotactemicity* of this tactophoneme, by which the number of all phonotactemes created out of this tactophoneme will be understood. Thus, the phonotactemicity of a tactophoneme is identical with the cardinal number of its phonotactemic range.

Ex 3.5 The phonotactemicity of the Polish tactophoneme $\{A, K, R\}$ equals 8, and the phonotactemicity of the Chinese tactophoneme $\{A, A, B, F, N\}$ equals only 2.

In order to obtain a general idea on the tactophonemes and the phonotactemes, being encoded by these tactophonemes, we shall subsequently give some of the results of the calculations worked out by Wierzchoń utilizing the dictionaries referred to above.

According to Wierzchoń's findings the Polish dictionary SJP-D contains **124.857** graphotactemes (that is, dictionary word-entries). And, these graphotactemes have been encoded by means of **91.740** tactographemes. Although Wierzchoń did not convert the graphotactemes and tactographemes into phonotactemes and tactophonemes, respectively, we can derive from his calculations approximate information on the number of the elements of these latter two sets of objects. Both the difference between the number of graphotactemes and that of phonotactemes, and the difference between the number of tactographemes and that of tactophonemes should be rather small.

The interpretation of the results obtained from the calculations based on the Chinese dictionary is, for reasons already mentioned above, more difficult. In particular, it was not possible here to automatically separate the word-entries from the syntagm-entries. Therefore, what was done, amounts to the extraction of the tactographemes from all the entries contained in the dictionary. Nevertheless, we can suppose that the

tactographemes comprised of a relatively large number of elements encode syntagmic graphotactemes rather than word graphotactemes. Keeping all this in mind should help us to interpret the calculation results which say that in Chinese dictionary **55.699** graphotactemes are encoded by means of **49.997** tactographemes.

4. Tactophonemic and phonotactemic dispersion of phones

A phone is almost always an element of more than just one tactophoneme, and is even more likely to be a member of more than just one phonotacteme. In order to reflect upon the occurrence of phones in these two kinds of objects we shall resort to the concept of dispersion. Before this will be done let us reiterate that each polyphonous tactophoneme is comprised of different phones. In consequence, a phone cannot repeat itself within the same tactophoneme. However, one and the same phone may recur in the same word-phonotacteme.

The set of all tactophonemes to which a given phone belongs will be called the *tactophonemic dispersion* of this phone, and the number of these tactophonemes will be called the *tactophonemic dispersion number* of this phone.

Analogously to the tactophonemic dispersion of a phone and to the number of this dispersion we can also introduce the *phonotactemic dispersion* of this phone, and its *phonotactemic dispersion number*. By the former the set of all phonotactemes in which this phone occurs is understood, and by the latter – the number of these phonotactemes.

As can be rightly inferred, the tactophonemic dispersion number of a phone cannot be greater than its phonotactemic dispersion number. It would seem that the ratio between the former number and the latter is also worthy of consideration. This ratio seems to reflect a certain aspect of the economy of using a given phone in coding phonotactemes by the corresponding tactophonemes. Thus, the smaller this ratio, the more economically a given phone is used in the coding referred to. The least value this economy attains in the case, if this ratio equals 1.

Phones of a given language can be compared with respect to their tactophonemic and phonotactemic dispersion as well as with respect to the corresponding dispersion numbers. These latter comparisons will allow us to establish rankings of phones according to their decreasing or increasing dispersion numbers. Thus, to put it more plainly, such rankings would show the extent of the participation of particular phones in creating tactophonemes and phonotactemes.

5. Phonotactemic efficiency of tactophonemes

The concept of phonotactemic efficiency, or the efficiency of phonotactemic coding, may be applied to both:

- (i) particular tactophonemes, and
- (ii) the family of all tactophonemes.

By *phonotactemic efficiency* of a tactophoneme we shall understand the ratio between the phonotactemicity of this tactophoneme and its phonicity. For the sake of quick recollection let us repeat that the former is identical with the number of the phonotactemes resulting from all the permissible sequentializations of the tactophoneme in question, whereas the latter is identical with the number of phones belonging to it. The smaller this ratio, the smaller the efficiency of phonotactemic coding of the tactophoneme in question is. In fact, this kind of efficiency merely says how many phonotactemes fall on average to each of the phones of the corresponding tactophoneme.

Ex 5.1 Let us compare two Polish tactophonemes $(A, K, R\}$ and $(A, K, R, T\}$ for their efficiency in question. The former tactophoneme encodes 8, and the latter 18 phonotactemes. Consequently, the efficiency of the former equals 2,6, and that of the latter 4,5.

Ex 5.2 Let us now compare three Chinese tactophonemes $(A, N\}$, $(A, D, N\}$ and $(A, F, U, Y\}$ for their efficiency. The first of these tactophonemes encodes the following three phonotactemes: *AN* (case, law case, record), *NA* (press down, restrain), and *NAN* (calamity, disaster). The second tactophoneme encodes only one phonotacteme *DAN*, but it represents various homonymous words with meanings such as: (thin, light); (dawn daybreak); (but, yet); (egg), etc. And the third tactophoneme encodes two phonotactemes: *FAYU* (French) and *YUFA* (grammar). Thus, the phonotactemic efficiency of these tactophonemes equals 3/2, 1/3, and 1/2, respectively.

An inquiry into Chinese phonotactics makes us aware of the low phonotactemic efficiency of the tactophonemes in this language, and at the same time of a great extent of a phenomenon which could be tentatively called ‘phonotactemic homonymization’ by which we mean the capability of one and the same phonotacteme to function as the representations for homonymous words. The homonymic efficiency of Chinese phonotactemes is amazing.

Phonotactemic efficiency, if extended to the whole family of tactophonemes, may be conceived of, as it seems, in various ways. The simplest approach to it suggests grasping it as the ratio between the number of all phonotactemes and the number of all tactophonemes. Defined in this way, phonotactemic efficiency indicates how many phonotactemes fall on average to each of the tactophonemes.

Ex 5.3 The 124.857 graphotactemes collected in the Polish dictionary (SJP-D) are encoded by means of 91.740 tactographemes. The resulting phonotactemic efficiency of the family of tactographemes amounts here thus to:

$$124.857 : 91.740 = \mathbf{1,36}$$

Ex 5.4 In the MDBG Chinese-English dictionary 55.699 Chinese grapho-tactemes are encoded by means of 49.997 tactographemes. The phonotactemic efficiency in

question amounts thus to:

$$55.699 : 49.997 = \mathbf{1, 11}$$

Although the values of efficiency given in these two examples do not concern phonotactemes or tactophonemes but graphotactemes and tactographemes, they give us a worthy insight into the efficiency which we are primarily interested in, since there is a close correlation between the phonic and graphic objects at issue.

6. A concept of tactophonome

Besides the concept of tactophoneme, defined as a set of phones capable of creating a phonotacteme, we shall also operate with the concept of tactophonome, which is expedient for our subsequent inquiry. In order to introduce it we shall avail ourselves of the *relation of tactophonemic equiphonicity*, which binds every two tactophonemes that are comprised of the same number of phones, that is, which are equiphonous.

Ex 6.1 The following two Chinese tactophonemes ([p], [rj], [i]) and ([t], [rj], [a]) are equiphonous. Both are triphonous (cf. *bing* ‘to be ill, illness’, *tang* ‘to lie’).

As an equivalence, the relation of tactophonemic equiphonicity specifies the corresponding classification of the family of tactophonemes. More precisely, this classification is thus the family of equivalence classes determined by this relation in the family of tactophonemes, and it will be called the *family of tactophonomes*. Each element of this family, that is, each tactophonome, emerges as the set of all tactophonemes having identical phonicity with a given tactophoneme. Or, to put it yet differently, each tactophonome results by gathering all those tactophonemes which are comprised of the same number of phones into one set.

With the family of tactophonomes at our disposal we can extend the concepts of phonicity, phonotactemic range, and phonotactemicity from tactophonemes to tactophonomes. The *phonicity of a tactophonome* will be conceived of as identical with the phonicity of each of its tactophonemes. Consequently, a tactophonome may be monophonous, biphonous, triphonous, etc.

By the *phonotactemic range* of a tactophonome the set of all phono-tactemes resulting from the sequentializations of the particular tactophonemes belonging to this tactophonome will be understood. And, the number of all these phonotactemes will be termed the phonotactemic onus (load) of this tacto-phonome or, simply, its *phonotactemicity*. Thus, the phonotactemicity of a given tactophonome is only the cardinal number of its phonotactemic range.

Particular tactophonomes may be comprised of various number of tactophonemes, that is, they may be mono-, bi-, tritactophonemeous, etc. By the tactophonemic power of a tactophonome or, simply, by its *tactophonemicity* the number of its tactophonemes will be understood.

7. Dependency between the phonicity of tactophonemes and their tactophonemicity and phonotactemicity

The number of tactophonemes as well as the number of the corresponding phonotactemes are dependent upon the number of phones of which they are comprised. In order to visualize these dependences we could plot tactophonomic phonicity, as an independent variable, on the x-axis, and tactophonomic tactophonemicity and tactophonomic phonotactemicity, as dependent variables, on the .y-axis. The resulting curves would graphically represent the dependences in question, that is, they would show how the number of tactophonemes as well as the number of phonotactemes change along with the change of the number of phones of which these tactophonemes and phonotactemes are created. In other words, we could see how many monophonous, biphonous, triphonous, etc. tactophonemes as well as the corresponding phonotactemes there are in a given language.

However, for lack of appropriate phonotactic data, which has already been pointed to above, we shall avail ourselves of the graphotactic data obtained for Polish and Chinese by Wierzchoń, who made his calculations based on the dictionaries referred to. His results for these dictionaries will be presented below in four tables. Table 7.1 (for Polish) and Table 7.3 (for Chinese) consist of three columns each A, B, and C, giving the following information:

- (i) in column A – the number of letter-types, that is, graphs,
- (ii) in column B – the number of tactographemes created out of the corresponding number of graphs, and
- (iii) in column C – the corresponding number of graphotactemes.

Thus, in column A the graphicity of a tactographeme, in column B its corresponding tactographemicity, and in column C its corresponding graphotactemicity are given. What is more, the tactographomic graphicity is given according to its increasing values.

The magnitudes listed in Table 7.1 and Table 7.3 are represented in Table 7.2 and Table 7.4, respectively, by plotting the values given in column A on the x-axis, and those given in B and C on the .y-axis, in systems of coordinates. The curves being obtained show the dependency of the number of tactographemes and graphotactemes upon the number of graphs (letter-graphs) out of which they are created.

The curves obtained in the above diagrams could be called tactographemicity and graphotactemicity curves, respectively. For both Polish and Chinese they assume the shape of a Gaussian curve. The number of tactographemes and graphotactemes gradually increases with the increase of the number of graphs (letter-types):

- (i) in Polish up to 8, and
- (ii) in Chinese up to 7,

and then gradually decreases. Thus, low and high graphicity (the number of graphs) is not favorable for high tactographemicity or high graphotactemicity.

Table 7.1. Polish (SJP-D)

A (GRAPHICITY)	B (TACTOGRAPHEMICITY)	C (GRAPHOTACTEMICITY)
1	32	32
2	87	131
3	681	1 093
4	2 352	3 656
5	5 638	8 574
6	10 636	15 789
7	15 770	22 700
8	18 360	25 494
9	16 233	21 285
10	11 391	13 991
11	6 281	7 379
12	2 870	3 204
13	1 061	1 170
14	274	283
15	61	63
16	13	13
	91 740	124 857

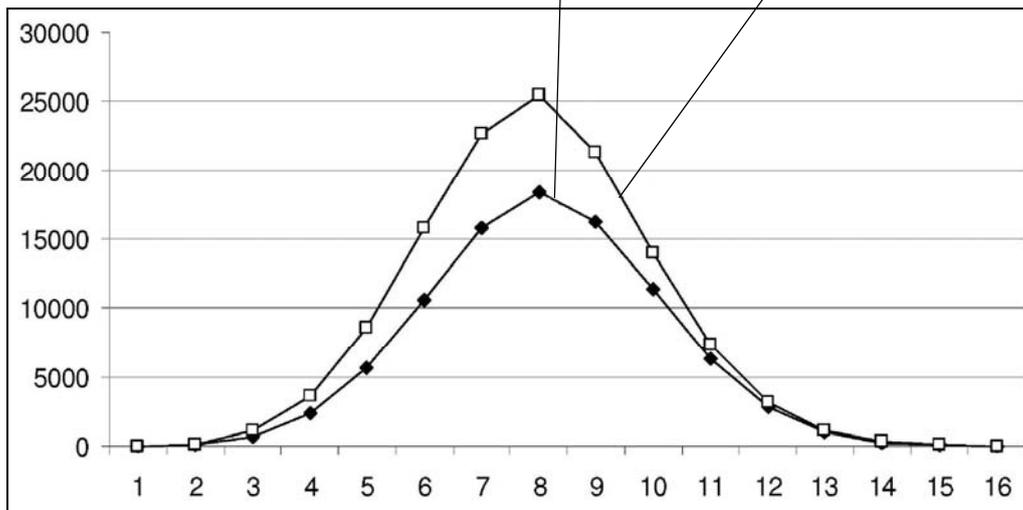
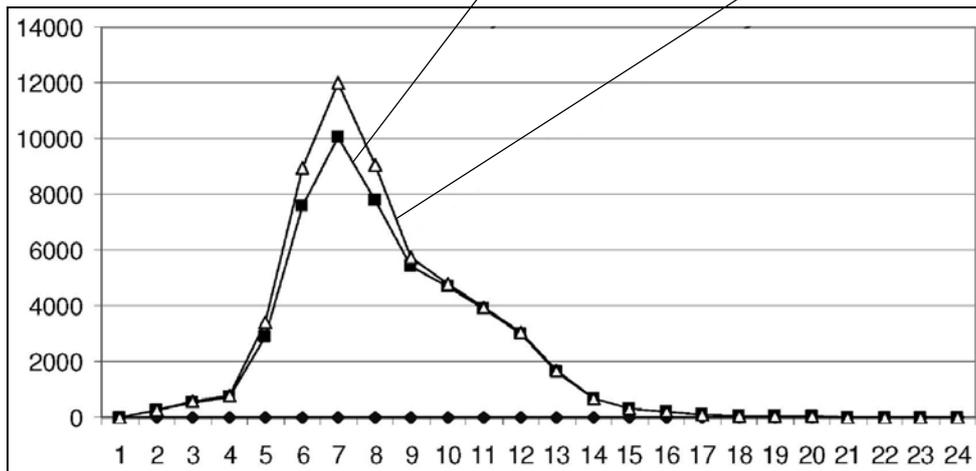
Table 7.2.

Table 7.3. Chinese (CC-CEDICT)

A (GRAPHICITY)	B (TACTOGRAPHEMICITY)	C (GRAPHOTACTEMICITY)
1	1	12
2	254	265
3	548	604
4	719	783
5	2 911	3 401
6	7 560	8 965
7	10 043	12 013
8	7 801	9 030
9	5 439	5 742
10	4 692	4 770
11	3 874	3 927
12	2 995	3 028
13	1 645	1 666
14	665	669
15	335	338
16	215	215
17	102	103
18	60	60
19	48	48
20	31	31
21	15	15
22	5	6
23	5	5
24	3	3
	49 997	55 699

Table 7.4.



Although these graphotactic results give us only an approximate insight into the corresponding dependences, we can nevertheless safely infer that the tactophonemicity and phonotactemicity curves for Polish and Chinese will behave similarly to the tactographemicity, and graphotactemicity curves, that is, that they will also assume the shapes of a Gaussian curve.

8. Phonotactemic efficiency of tactophonemes

In section 5 we have defined the phonotactemic efficiency of tactophonemes. Now, with the tactophonemes at our disposal, we can extend this kind of efficiency also to these objects.

The *phonotactemic efficiency of a tactophoneme* will be conceived of as the ratio between its phonotactemicity and tactophonemicity. The smaller this ratio, the smaller the efficiency of the phonotactemic coding of the tactophoneme in question is.

The comparison of all tactophonemes for their phonotactemic efficiency should reveal which tactophonemes are more efficient than others, that is, which tactophonemic phonicities are conducive to efficiency being considered.

9. Tactophonomic phone-basis

As should be remembered all the tactophonemes belonging to a given tactophoneme are equinumerous. We can now ask the question which phones form all and only those tactophonemes. And, the set of all these phones will be called the *phone-basis* for the considered tactophoneme. Formally, the phone-basis for this tactophoneme is obtained from the set-theoretical summation of all its tactophonemes. And, what is more, each tactophoneme is associable with exactly one corresponding phone-basis.

In addition to the tactophonomic phone-basis we shall introduce the concept of *tactophonomic phone-basis phonicity*, which will be understood as the number of all phones which belong to the tactophonomic phone-basis. Thus tactophonomic phonicity and tactophonomic phone-basis phonicity are two different concepts.

Having distinguished the concept of phone-basis we can ask the following questions:

- (i) How the particular tactophonomic phone-bases are related to the family of all phones, that is, are they smaller than this family or exhaust it completely.
- (ii) What is the phonicity of the particular tactophonomic phone-bases, that is, the number of their phones.
- (iii) How the phonicity of a tactophonomic phone-basis is dependent upon the tactophonomic phonicity.

10. Tactophonomic equiphony and disphony

Tactophonemes may be similar and dissimilar with respect to the phones of which they are comprised. In other words, the phones which belong to different tactophonemes may be identical or different. We shall approach certain aspects of this kind of phonetic similarity/dissimilarity in terms of equiphony and disphony, while operating with the concepts of equiphonous and disphonous bases, which, in turn, will be subdivided into broad and narrow ones, respectively. All four kinds of resulting basis will refer to each particular tactophoneme.

The *broad equiphonous basis* for a tactophoneme is comprised of all those phones each of which is an element of at least two tactophonemes belonging to this tactophoneme. And, the *narrow equiphonous basis* for a tactophoneme is comprised of all those phones each of which is an element of every tactophoneme belonging to this tactophoneme. As is easily noticeable, the latter of these bases is only a particular kind of the former. Some tactophonemes may have one or even both of these bases empty.

The *broad disphonous basis* for a tactophoneme is comprised of all those phones each one of which is an element of the phone-basis of this tactophoneme but is not an element of every tactophoneme belonging to it. And, the *narrow disphonous basis* for a tactophoneme is comprised of all those phones each of which is an element of the phone-basis of this tactophoneme but is not an element common to some two tactophonemes belonging to it.

As can be rightly inferred, the narrow disphonous basis will always be included in the broad one, which, in turn, is non-empty for every polytactophonemic tactophoneme, since any of its two elements must differ at least with regard to one phone.

11. Concluding remarks

A fragment of the phonotactological theory sketched out above should be viewed only as a preliminary draft aiming at the identification of the areas of phonotactics where it could be effectively applied rather than at offering already satisfactory results. More precisely, it was the author's intention to examine whether the concepts with which this theory operates do adequately capture at least some relevant properties of the phonotactic structure of words. And, these concepts include: *phonotacteme*, *tactophoneme*, *tactophonemic* and *phonotactemic dispersion of phones*, *phonotactemic efficiency of tactophonemes*, *tactophoneme*, *tactophonomic phone-basis*, *equiphony*, *disphony*, to mention only some of them.

The author is also fully aware of the approximate nature of the exemplifications being given. The unavailability of suitable phonotactic language material certainly weakened the value of these exemplifications. But nevertheless the proposed theory may turn out to be a source of inspirations which may result in more adequate elaborations of general and particular phonotactology.

The author would also like to hope that the journey accomplished in the present, still imperfect, phonotactological vehicle into the enormous expanse of words will contribute to making at least one further small step towards a better understanding of the phonotactic reality of ethnic languages, a reality full of enigmas and surprises. However, if this hope is unfounded, that is, if the reader will get the impression of having wasted time on this article, then all that's left to do is to apologize for my misconceived approach to the reality in question.

Bibliography

- Bańczerowski, Jerzy. 1992. Is generative phonology consequent upon structural phonology. In: Fisiak, Jacek; Puppel, Stanisław (eds). 1992, 247–307.
- . 2009. The concept of tactophoneme (in press).
- Batóg, Tadeusz. 1967. *The axiomatic method in phonology*. London: Routledge & Kegan Paul.
- Chen, H.-C.; Tzeng, O.J.L. (eds). 1992. *Language processing in Chinese*. Amsterdam: North-Holland and Elsevier.
- Duanmu, San. 2007. *The phonology of standard Chinese*. Oxford: Oxford University Press.
- Fisiak, Jacek; Puppel, Stanisław (eds.). 1992. *Phonological investigations*. Amsterdam: John Benjamins.
- Hoosain, R. 1992. Psychological reality of the word in Chinese. In: Chen, H.-C.; Tzeng, O.J.L. (eds). 1992. 111–130.
- Jones, D. 1950. *The phoneme: Its nature and use*. Cambridge: Heffer.
- Li, Charles N.; Thompson, Sandra A. 1989. *Mandarin Chinese. A functional reference grammar*. Berkeley–Los Angeles–London: University of California Press.
- Packard, Jerome L. 2000. *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.
- . (ed.). 1997. *New approaches to Chinese word formation: Morphology, phonology and the lexicon in modern and ancient Chinese*. Trends in Linguistics Studies and Monographs 105. Berlin and New York: Mouton de Gruyter.
- Sproat, Richard; Shih, Chilin. 1996. A corpus-based analysis of Mandarin nominal root compound. *Journal of East Asian Linguistics* 5, 49–71.
- Wierchoń, Piotr. 2004. *Gramatyka diakrytologiczna. Studium ortograficzno-kwantytatywne*. Poznań: Wydawnictwo Naukowe UAM.
- Zabrocki, Ludwik. 1962. Phon, Phonem und distinktives Morphem. *Biuletyn Fonograficzny* V, 59–87.
- Zhou, X.; Murslen-Wilson, W. 1994. Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes* 9, 393–422.
- . 1995. Morphological structure in the Chinese mental lexicon. *Language and Cognitive Processes* 10, 545–600.