

## Statistical Analysis of Chinese Phonemic Contrast

Shuiyuan Yu<sup>a</sup> Chunshan Xu<sup>b</sup> Haitao Liu<sup>c</sup> Yudong Chen<sup>d</sup>

<sup>a</sup>School of Computer and <sup>d</sup>International College for Chinese Language, Communication University of China, Beijing, <sup>b</sup>Foreign Languages Department, Anhui University of Architecture, Hefei, and <sup>c</sup>Center for Language-Behavior Patterns, Zhejiang University, Hangzhou, P.R. China

### Abstract

Two phonemes that may induce minimal pairs constitute a *phonemic contrast*. Some phonemic contrasts may disappear for various reasons, which, nevertheless, does not seem to seriously impede linguistic communication. Does it mean that the disappeared phonemic contrasts are unimportant? In our study, we calculated the proportions (here termed degree of contrast) of minimal pairs to the words in which the two contrastive phonemes occur and explored the role of phonemic contrasts in the phonemic combinations. The degree of contrast of phonemes reflects the relation between phonemes. Our results indicate that (1) the average degree of contrast of Chinese phonemes declines exponentially with the increase in the number of syllables, rapidly approaching zero; (2) the average degree of contrast of Chinese consonants that differ from each other in only one distinctive feature and of the consonants that are absent in some Chinese dialects is significantly higher than that of other consonants; (3) the degree of contrast of Chinese consonants that differ from each other in only one distinctive feature is not significantly different from that of the consonants absent in some Chinese dialects; (4) Chinese phonemic combinations exhibit high degree of sparsity, which increases exponentially with the number of syllables and rapidly approaches 1. All these results show that the high degree of sparsity and the low degree of contrast of human languages not only leave enough room for new words, new dialects and new languages to appear but also contribute to effective and reliable communication, because a few phonemic mistakes are not likely to cause wrong decoding (sound recognition) and failed communication.

Copyright © 2012 S. Karger AG, Basel

### 1. Introduction

Phonemes are vital for human languages because they are the smallest units that can differentiate meaning. However, some speakers who are unable to articulate some

**KARGER**

Fax +41 61 306 12 34  
E-Mail [karger@karger.ch](mailto:karger@karger.ch)  
[www.karger.com](http://www.karger.com)

© 2012 S. Karger AG, Basel  
0031–8388/11/0684–0201  
\$38.00/0  
Accessible online at:  
[www.karger.com/pho](http://www.karger.com/pho)

Haitao Liu  
School of International Studies, Zhejiang University  
No. 866 Yuhangtang Road  
CN–310058 Hangzhou (P.R. China)  
Tel./Fax +86 571 88981923, E-Mail [lhtzju@gmail.com](mailto:lhtzju@gmail.com)

phonemic features can still perfectly communicate with others who can correctly articulate these features. Some physiological problems may result in slurred pronunciation and mother tongue interference may lead to second language learners' inability to distinguish and articulate some phonemes. Nevertheless, these problems usually bring no great troubles to communication with other people. Here naturally arise two questions: why communication is not hindered by the loss of some phonemes and the ensuing reduction in differentiability of the corresponding language, and whether these lost phonemes have some distinctive characteristics. Both questions are highly relevant to one important feature of human language – double articulation (also known as duality of patterning), which is now held to be one of the fundamental differences between human language and animal communication [Hockett, 1960; Jackendoff, 2002; Zuidema and de Boer, 2009].

A word is composed of phonemes and the replacement of any phoneme in this word will surely change it into another word or a meaningless phonemic sequence. The study of the relations among the phonemes, which can shed light on the ways phonemes are combined into words, is important in the research on the combinatorial property of language. Different phonemes can occur in the same environment, giving rise to minimal pairs, which is known as *phonemic contrast*. Contrast is important for a language system to communicate message. For two phonemes in a contrast, the more words they can distinguish, the more important they are in this language. Different means are employed in a language to convey information. Every means has a functional load, that is, the degree to which a certain language depends on it to communicate message. The notion of functional load, which dates back to the Prague School [Mathesius, 1929; Trubetzkoy, 1939], is important in linguistics and has a close relation with sound changes in a language system [King, 1967]. Phonologically, *functional load* [Hockett, 1955; King, 1967], which can be quantified by the number of words distinguished by two contrastive phonemes, may be seen as referring to the degree to which a language relies on this phonemic contrast. In other words, the functional load reflects the importance of the difference between two contrastive phonemes in a certain language.

Given that the functional load is a metric of the importance of a contrast, it seems reasonable to follow that a contrast with lower functional load (less important in a certain language) will be more likely to disappear in the process of evolution – the so-called Martinet hypothesis [King, 1967; Martinet, 1933, 1955; Peeters, 1992]. Martinet [1933, 1955] holds that the communication function of language prevents the merger of phoneme contrasts of high functional load (though it does not necessarily mean that the contrasts with low load are bound to disappear). It can be inferred from Martinet's hypothesis that the relation between the functional load and sound change is as follows: 'Given the assumption that communication must be maintained, and given the fact that sounds change, it simply seems indisputably true that sounds should change in a way which does the least damage to communication' [King, 1967]. Nonetheless, this hypothesis is questioned by some scholars [Hoenigswald, 1960] because Martinet failed to provide any empirical evidence. What is more, some phonemes may undergo phonetic and phonemic changes simply owing to their 'intrinsic' features, or rather, the articulatory, acoustic and aerodynamic properties of their phonetic realizations [Ohala, 1981, 1997].

Research into the evolution of German and Cantonese has also revealed that the contrasts lost in evolution are not necessarily those with low functional load [Surendran and Niyogi, 2006]. Though not enough to disprove Martinet's conjecture, this discovery at least points to the fact that the disappearance of phoneme contrasts can hardly be

solely attributed to their low functional load – contrasts with high functional load may also disappear in the evolution of a language.

The hypothesized disappearance of contrast with low load implies that, after a long process of evolution, all the remaining phonemic contrasts in a language will be of considerable importance. In that case, it is highly probable that the replacement or removal of any phoneme in a word will produce another existing word. Thus, with all possible phonemic combinations in this language actually employed and realized, the highest coding efficiency is achieved in this language. In this way, the phonemic system is optimized because as many words as possible are constituted with as few phonemes as possible, which conduces to the economy of not only storage but motor and perceptual representations. Nevertheless, for a communication system, high efficiency means, according to information theory, low reliability. In such a language system, the mistake or missing of one single phoneme will lead to the misinterpretation of the entire word into another. Therefore, such a language system puts high demand on articulation, sound discrimination and transmission environment, rendering online processing extremely difficult. In short, the relation between functional load and sound change is probably not as monotonous as claimed by Martinet [1933, 1955].

Apart from this defect, the notion of functional load has been lacking a clear definition. Meyerstein [1970] noticed the great difficulty in giving an exact definition of functional load, especially an operationalized one. Presently, functional load is generally held to be the differential ability of phonemic contrasts in a certain language. Hockett [1955] derived the first formula to compute functional load, which is based on the assumption that functional load is quantitatively equal to information loss. Later on, other researchers presented more general formulae [Carter, 1987; Ingram, 1989; King, 1967; Surendran and Niyogi, 2006]. Though varying in computing methods and applicable scopes, all these formulae are invariably based on the multiplication of the frequency by the number of words (minimal pairs) differentiated by a contrast. As a result, the functional load of some contrasts may be considerably high because of the high frequency of some words, and vice versa. However, a language changes all the time and so do the frequencies of its words. As a result, functional load may also change with the variation of word frequency. Thus, it is improper to solely ascribe sound changes to functional load, not to mention the complex relations between sound changes and many other factors such as society, history, geography etc. [Atkinson, 2011; Hay and Bauer, 2007; Manica et al., 2007]. In all these formulae, the functional load is the multiplication of number and frequency, failing to clearly report either the number of minimal pairs or their frequency. Therefore, separately dealing with these two factors may help us get a clear picture of the contrasts lost in the process of evolution.

The number of minimal pairs resulting from a phonemic contrast reflects the relation between the contrastive phonemes and reveals the role of the contrast in a language system rather than in language use. Absence of minimal pairs means that there is a complementary relation between the two phonemes, and a large number of minimal pairs imply that there must be a high degree of contrast (DC) between them. In fact, a language usually has enough phonemes to generate much more minimal pairs and possible phonemic combinations than what is required to form all the words in this language, a mechanism both increasing the redundancy of the language system and enhancing the reliability of communication. In this sense, the distribution of the numbers of minimal pairs may be a clue to how efficiency and reliability are balanced in a language.

The study of the relation among phonemes also contributes significantly to the design and development of feature-based speech recognition systems. Presently, the mainstream automatic speech recognition systems based on phonetic identities of time frames have failed to work satisfactorily in real environments outside of the laboratory. By computing the phonetic identities of temporally continuous data frames, these systems establish trajectories in phonetic state space. Different trajectories represent different utterances. The major defect of this approach is that the missing or distortion of one single data frame, which may be caused by noise interference or wrong articulation, can take the trajectory to a different direction, resulting in a totally wrong recognition. This approach is different from human auditory mechanisms, in which phonetic features and the time relation between them are usually believed to be the foundation of speech perception [Esser et al., 1997; Stevens, 1998, 2002; Teich and Lowen, 1994]. This means that a mathematical model based on distribution properties of phonemic features may be of considerable importance to the design of the speech recognition systems based on the detection of phonemic features. Therefore, it is necessary to explore the relations between phonemes.

The second section of this paper is a brief introduction of the corpora used in this study. The third section reports the DCs of different types in these corpora and the degrees of sparsity for words of different lengths. The fourth section presents the conclusion and the discussion.

## 2. Materials

The corpora used in our study include three genres: news, stories and daily conversations. The themes of the stories cover ancient legends, history, literature, modern sciences, education, etc. The news was collected in 2007 and 2008 from a famous Chinese website (<http://news.sogou.com/>). The daily conversations, consisting of roughly 840 sentences, are all taken from textbooks of oral Chinese ([http://www.china.org.cn/learning\\_chinese/dialogue/](http://www.china.org.cn/learning_chinese/dialogue/)).

We have extracted, from these corpora, 90,000 common Mandarin Chinese words. Of these words, 5,634 (6.39%) are monosyllabic, 47,822 (54.24%) are bisyllabic, 21,916 (24.86%) are trisyllabic and 11,882 (13.48%) are quadrisyllabic. Words with five or more syllables only account for 1.03% of all words. In our study words and compounds are not strictly distinguished.

We extracted words from corpora because, on the one hand, it is unfeasible to collect all existing Mandarin Chinese words and, on the other hand, most of existing Mandarin Chinese words are professional terms and uncommon words which are used quite rarely. In comparison, words extracted from real corpora can better reflect the synchronic state of the language in the process of evolution. In fact, this set of extracted words, supplemented with only a few proper names, well supported the automatic Chinese word segmentation, which is the foundation of some previous works [Yu et al., 2011]. This indicates that words extracted from corpora can faithfully represent the Mandarin Chinese vocabulary.

Our study excluded words with five or more syllables because, on the one hand, being mostly names of institutions, nations and places, those items are, strictly speaking, not words, and on the other hand, they only account for a very tiny part (about 1%) of the whole set of words.

## 3. The Metric of the DC of Phonemes

We define the DCs of two phonemes contrasting with each other as the proportions of minimal pairs to the words in which these two phonemes can occur. Since the two contrasting phonemes can occur in different numbers of words, the minimal pairs may

account for different portions of the words in which each phoneme occurs, or, the two phonemes in a contrast may have different DCs. For example, in language L, /a/ and /a'/ are contrastive phonemes; /a/ occurs in /ax/, /ay/, /az/, /au/, while /a'/ in /a'w/, /a'u/ and /a'x/. There are two minimal pairs: [/ax/, /a'x/] and [/au/, /a'u/], which account for half of the words in which /a/ occurs and two thirds of the words in which /a'/ occurs. It is therefore of considerable importance to separately calculate for each phoneme the proportion of minimal pairs to the words in which it occurs, for this method can help us obtain the accurate distribution of phonemic contrasts. We define DC as follows:

For /a/ and /a'/ in a phonemic contrast, the DC of /a'/ is:

$$C(a'/a) = \frac{\text{the number of minimal pairs of / a / and / a' /}}{\text{the number of words in which / a' / occurs}} \quad (1)$$

The DC has the following properties: (i) the DC of /a'/ is usually not equal to that of /a/, i.e.,  $C(a'/a) \neq C(a/a')$ ; (ii) the value of the degree ranges from 0 to 1, i.e.,  $0 \leq C(a/a') \leq 1$ . Value 0 means the two contrastive phonemes are in complementary distribution.

According to the formula proposed, between DC and functional load is the following relation:

$$FL = \sum_{i=1}^N p_i$$

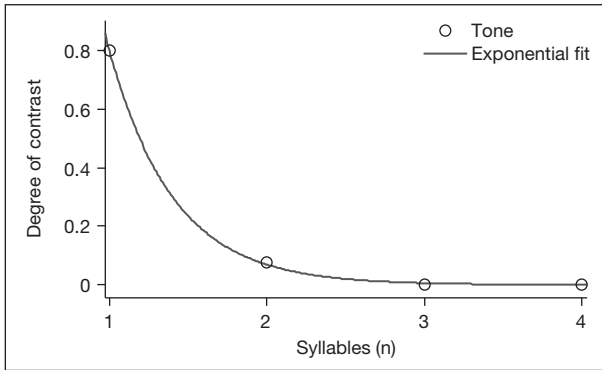
where FL is functional load,  $p_i$  is the frequency of the  $i$ -th minimal pair of these two phonemes and  $N$  is the number of minimal pairs, i.e.,  $N = (\text{the number of words in which /a'/ occurs}) \times C(a'/a) = (\text{the number of words in which /a/ occurs}) \times C(a/a')$ . The most important difference between these formulae proposed in the literature and formula (1) that we defined is that the latter eliminates the impact of the frequency of minimal pairs. The advantage of formula (1) is that the calculation takes into account only linguistic rules, disregarding the frequencies of words and phonemes.

### 3.1. Measuring the DC

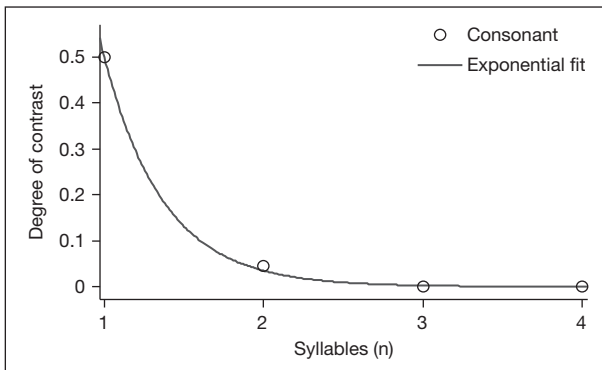
We calculated, with formula (1), the DC of phonemes of different types (vowels, consonants, tones) in words of different lengths. Figure 1 shows the relation between word length (in terms of the number of syllables) and the average DC of tones; the fitting function is  $f(x) = 9.256 \cdot \exp(-2.449 \cdot x)$  and the goodness of fitting is adjusted  $R^2 = 0.9997$ .

The phonological rules of Mandarin Chinese confine most Mandarin Chinese consonants to the beginning position of a syllable and the nasal /ŋ/ to the coda of a syllable. Nasal /n/ can occur at both the beginning and coda of a syllable. In other words, no consonant except nasal /n/ can form minimal pairs with /ŋ/. So we simply calculated the DCs of consonants occurring at the beginning of syllables, leaving nasal /ŋ/ out. Figure 2 shows the relation between word length and the average DC of consonants; the fitting function is  $f(x) = 7.249 \cdot \exp(-2.673 \cdot x)$  and the goodness of fitting is adjusted  $R^2 = 0.9991$ .

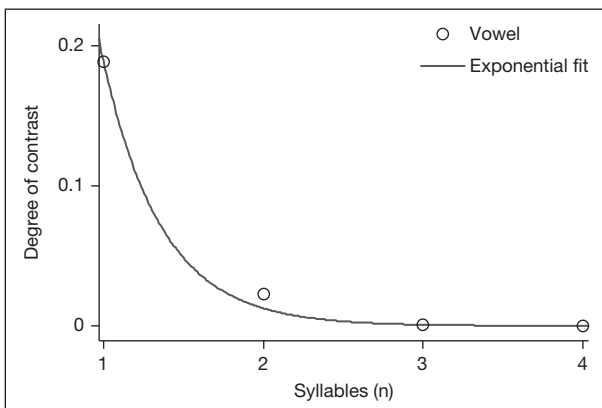
Figure 3 shows the relation between the word length and the average DC of vowels; the fitting function is  $f(x) = 2.864 \cdot \exp(-2.716 \cdot x)$  and goodness of fitting is Adjusted  $R^2 = 0.9939$ . These three figures indicate that the average DC decreases exponentially with the increase in word length, rapidly approaching zero.



**Fig. 1.** Average DC of tones as a function of word length (by number of syllables).



**Fig. 2.** Average DC of consonants as a function of word length.



**Fig. 3.** Average DC of vowels as a function of word length.

DC reflects the relation between phonemes. If two phonemes are in complementary distribution, the DC between them is 0; if they give rise to many minimal pairs, the DC between them is rather high. To find out the phonological significance of the DC as calculated according to our formula, we conducted a clustering analysis into the DCs of contrastive consonants in monosyllables.

The clustering analysis is conducted on the basis of DC, which is taken as the distance between phonemes: the phonemes with smaller distances between them are clustered into one group. In table 1, phonemes with the same number belong to the same group and, as can be seen, the consonants are therefore clustered into three groups. As can be seen, the clustering hinges on /tɛ/, /tɛ'/, /ɛ/, retroflex, and /ts/, /ts'/, /s/, apical, which reflects the combinational relations between them and the rest of the syllables (mainly vowels) where they appear. In Mandarin Chinese, a syllable is composed of the initial (the consonant occurring at the beginning of the syllable) and the final (vowels and nasals that constitute the rest of the syllable). The collocation relations between initials and finals usually consist of the place of articulation of initials and the manner of articulation of beginning vowels of finals. If initials have the same place of articulation and the beginning vowels of finals have the same manner of articulation, the collocation relations between these initials and finals are generally the same. The clustering result as shown in table 1 also shows this tendency. In other words, the DC that we defined reflects the relation between phonemes. For example, it can be seen in table 1 that there are complementary relations between consonants /ts/, /ts'/, /s/, /tɕ /, /tɕ'/ and /ɕ/. In other words, as far as monosyllables are concerned, the DC conforms well to the combinatorial rules of phonemes. However, such conformity declines with an increasing number of syllables. Phonologically, the phonemic contrasts induced by a single distinctive feature reflect a rather important phonological relation, which gives human phonemes a symmetry property. The following section is concerned with the characteristics of the degree of such phonemic contrasts.

### 3.2. DC of Contrastive Consonants Differing in Only One Distinctive Feature

A phoneme can be described in terms of a group of distinctive features. Two contrastive phonemes may differ in only one distinctive feature, implying a somewhat symmetrical distribution of phonemes. Here, the symmetric distribution refers to the phenomenon that, in a language, a phoneme with a distinctive feature +A is usually accompanied by another phoneme which is identical except for a distinctive feature –A. For example, most languages have both /p/ and /b/, which are the same except for one distinctive feature – voicedness. Of course, there are exceptions: there is only /b/ but not /p/ in Arabic. We can describe phonemes with a matrix of distinctive features and compare the DC of contrastive consonants differing in only one distinctive feature with that of other consonants so as to find out whether the former have some special properties.

A Chinese dialect may lack some Mandarin Chinese phonemic contrasts. As a result, when speaking Mandarin Chinese, the dialect speakers are likely to find it quite difficult to articulate certain phonemic contrasts, which, however, does no serious harm to their communication. Given that the degree of phonemic contrast reflects the collocation relation between phonemes, we wanted to find out whether there is any difference in DC between the phonemes that are absent in some dialects and the phonemes that contrast with respect to only one distinctive feature.

The Mandarin Chinese consonants absent in some Chinese dialects include /n/ and /l/, /f/ and /x/, /ts'/ and /tɕ'/, /s/ and /ɕ/, and /ts/ and /tɕ/. The contrastive consonants

**Table 1.** Clustering of DC of consonants in monosyllables

	b	ts'	t	f	k	x	te	k'	l	m	n	b'	te'	ʒ	s	t'	ɛ	ts	tʂ'	ʃ	tʂ	
b		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
ts'	3		3	3	3	3	1	3	3	3	3	3	1	3	3	3	1	3	3	3	3	3
t	3	3		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
f	3	3	3		3	3	2	3	3	3	3	3	2	3	3	3	2	3	3	3	3	3
k	3	3	3	3		3	2	3	3	3	3	3	2	3	3	3	2	3	3	3	3	3
x	3	3	3	3	3		2	3	3	3	3	3	2	3	3	3	2	3	3	3	3	3
te	3	1	3	2	2	2		2	3	3	3	3	3	2	1	3	3	1	1	1	1	1
k'	3	3	3	3	3	3	2		3	3	3	3	2	3	3	3	2	3	3	3	3	3
l	3	3	3	3	3	3	3	3		3	3	3	3	3	3	3	3	3	3	3	3	3
m	3	3	3	3	3	3	3	3	3		3	3	3	3	3	3	3	3	3	3	3	3
n	3	3	3	3	3	3	3	3	3	3		3	3	3	3	3	3	3	3	3	3	3
b'	3	3	3	3	3	3	3	3	3	3	3		3	3	3	3	3	3	3	3	3	3
te'	3	1	3	2	2	2	3	2	3	3	3	3		2	1	3	3	1	1	1	1	1
ʒ	3	3	3	3	3	3	2	3	3	3	3	3	2		3	3	2	3	3	3	3	3
s	3	3	3	3	3	3	1	3	3	3	3	3	1	3		3	1	3	3	3	3	3
t'	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3		3	3	3	3	3	3
ɛ	3	1	3	2	2	2	3	2	3	3	3	3	3	2	1	3		1	1	1	1	1
ts	3	3	3	3	3	3	1	3	3	3	3	3	1	3	3	3	1		3	3	3	3
tʂ'	3	3	3	3	3	3	1	3	3	3	3	3	1	3	3	3	1	3		3	3	3
ʃ	3	3	3	3	3	3	1	3	3	3	3	3	1	3	3	3	1	3	3		3	3
tʂ	3	3	3	3	3	3	1	3	3	3	3	3	1	3	3	3	1	3	3	3		3

differing in one distinctive feature include /p/ and /p'/, /t/ and /t'/, /k/ and /k'/, /ts/ and /ts'/, /tʂ/ and /tʂ'/, and /te/ and /te'/. The comparison is based on the Kolmogorov-Smirnov test of their average DCs. For each contrast, there are two DCs: C(a/a) and C(a/a'). The result of Kolmogorov-Smirnov test is shown in table 2. All p values are larger than 0.05, therefore the null hypothesis that the two mean degrees are the same cannot be rejected. That is, there is no significant difference between the two means.

Then we conducted another Kolmogorov-Smirnov test to make certain whether the DC of the consonants contrasting with each other in one feature is different from that of other consonants. The result is shown in table 3. In monosyllables, bisyllables and quadrisyllables, the consonants contrasting with each other with respect to one feature, according to table 3, have significantly higher DC than other consonants, but, as can be seen in table 2, roughly the same DC as the consonants absent in some Chinese dialects. In other words, we did not observe the assumed dependence of the absence of some phonemic contrasts on their hypothesized low DC in Mandarin Chinese. In fact, the DC of those contrastive consonants is significantly higher than that of consonants that contrast with each other in more than one distinctive feature.

Although the consonants absent in some dialects exhibit no lower DC than other consonants, their absence can possibly be explained in terms of the small number of words they help to form. Therefore we calculated the frequencies of these consonants



**Table 2.** Results of Kolmogorov-Smirnov tests of average DC of absent consonants and average DC of contrastive consonants differing in one distinctive feature

	Monosyllable	Bisyllable	Trisyllable	Quadrisyllable
Mean DC of contrastive consonants differing in one distinctive feature	0.7932	0.0612	0.0018	0.0003
Mean DC of absent consonants	0.7035	0.0629	0.0016	0.0007
Rejection of the equality between the two means	0	0	0	0
Significance level (p value)	0.0873	0.9304	0.6259	0.1307

**Table 3.** Results of Kolmogorov-Smirnov tests of average DC of contrastive consonants differing in one distinctive feature and average degree of other consonants

	Monosyllables	Bisyllables	Trisyllables	Quadrisyllables
Mean DC of contrastive consonants differing in one distinctive feature	0.7621	0.0632	0.0018	0.0005
Mean DC of other consonants	0.4867	0.0441	0.0012	0.0001
Rejection of the equality between the two means	1	1	0	1
Significance level (p value)	0	0.0020	0.6072	0.0055

**Table 4.** Frequency of occurrence: consonants (%)

	b	ts'	t	f	k	x	te	k'	l	m	n	b'	te'	ʒ	s	t'	ε	ts	tʂ'	ʂ	tʂ
Mono-syllable	5.00	3.88	5.52	2.59	4.92	5.69	4.23	4.23	7.51	5.09	5.09	5.18	4.40	2.59	3.54	5.95	4.40	4.14	5.69	5.00	5.35
Bi-syllable	5.37	2.32	6.79	4.11	5.39	6.23	8.44	2.77	6.14	4.38	2.18	2.75	4.50	2.02	2.40	5.09	7.00	3.26	4.82	7.42	6.62
Tri-syllable	5.38	1.82	6.70	3.83	4.95	6.58	8.58	2.12	7.19	3.55	2.12	2.46	4.85	1.68	2.45	4.48	9.71	2.91	3.93	8.47	6.23
Quadri-syllable	6.68	2.03	6.86	3.80	5.16	5.54	8.28	2.52	7.11	4.47	2.05	1.76	4.94	3.32	2.95	4.23	6.99	3.09	3.58	7.86	6.79

in words. Table 4 shows the frequencies of Chinese consonants (only the consonants that can occur at the beginning of a word) in words of different length, which can be seen as the metric of the importance of these consonants.

Obviously, absent consonants such as /x/, /n/, /tʂ/, /tʂ'/, /ʂ/ (implying the absence of phonemic contrasts like /f/-/x/, /l/-/n/, /ts/-/tʂ/, /ts'/-/tʂ'/, /s/-/ʂ/) are not the least frequent ones in Mandarin Chinese. Actually, these absent consonants occur even more frequently than some consonants that are not absent.

Low DC implies a high degree of sparsity, that is, many possible phonemic combinations are not actually realized and used in this language. The degree of sparsity

also reflects indirectly the reliability of the language system. The following section is devoted to the degree of sparsity of phonemic combinations.

### 3.3. Degree of Sparsity of Phonemic Combinations

If in one language all possible phonemic combinations allowed by phonological rules are realized as existent syllables and all allowed multisyllables are realized as existent words, there is, in terms of phonemic combinations, zero redundancy in this language, that is, the highest efficiency is achieved when phonemes in this language combine to produce words. Nevertheless, information theory tells us that a language (as a communication system) with such high efficiency (i.e., zero redundancy) is rather unreliable. We can calculate the degree of sparsity of phonemic combinations to measure efficiency. The formula is as follows:

$$s = 1 - \frac{\text{the amount of actually realized combinations}}{\text{the amount of all legitimate combinations}} \quad (2)$$

According to this formula, the higher the degree of sparsity, the fewer legitimate phonemic combinations are actually realized and used as syllables and words.

Mandarin Chinese phonological rules prescribe that a Chinese monosyllable be composed of: one consonant (optional), one to three vowels, nasal /n/ or /ŋ/ (optional) and a suprasegmental feature – tone. [There are actually five tones in Mandarin Chinese, but the fifth ‘neutral’ tone (qing sheng), which is only used rarely in special cases, is excluded from our study.] In addition to this phonological rule, there are some other combinational constraints on the structure of a Mandarin Chinese syllable. For example, only /n/ and /l/ can immediately precede all vowels; /tɕ/, /tɕʰ/, /ɕ/ can only collocate with finals beginning with [i] or [y]. Duanmu [2009] proposed a method to calculate the number of all possible Chinese syllables: ‘A full syllable in Standard Chinese can have up to four underlying sounds, CGVX, where C is a consonant, G a glide, V a vowel, and X an off glide of a diphthong or a consonant. In addition, each full syllable can carry one of four tones. In principle, there are 1,900 possible full syllables without tonal contrasts, or 7,600 including tonal contrasts.’ Duanmu’s [2009] approach is based on existing Mandarin Chinese phonemic combinations and therefore is adopted in our study. According to our study, only 1,294 of all possible syllables are used, which means a 0.83 degree of sparsity ( $1 - 1,294/7,600 = 0.83$ ). Mandarin Chinese words are usually composed of more than one syllable. The degree of sparsity of bisyllables is 0.9992 ( $1 - 47,822/7,600^2 = 0.9992$ ); as to trisyllables and quadrisyllables, the degree is very close to 1 ( $1 - 21,916/7,600^3 \approx 1$ ;  $1 - 11,882/7,600^4 \approx 1$ ). Apparently, in Mandarin Chinese, possible phonemic combinations far outnumber those actually used. On the whole, the fitting function of the degree of sparsity is:  $f(x) = 1 - 0.1703 * x^{-13.04}$  (x can be 1, 2, 3, or 4, namely, the number of syllables); the goodness of fit is: adjusted  $R^2 = 1$ .

It is also reasonable to consider only the 1,294 actually used monosyllables that compose multisyllables. In this case, the degrees of sparsity of bisyllables, trisyllables and quadrisyllables are 0.9693 ( $1 - 47,822/1,294^2$ ), 0.9989 ( $1 - 21,916/1,294^3$ ) and roughly 1 ( $1 - 11,882/1,294^4$ ), respectively.

As can be seen, whichever method is used, the high degree of sparsity persists – only a small portion of theoretically possible combinations are used in Mandarin

**Table 5.** The t test of the DCs of phonemes in a phonemic contrast

		Monosyllable	Bisyllable	Trisyllable	Quadrisyllable
Vowel	reject same distribution	1	0	0	0
	significance level	0.0137	0.0589	0.0598	0.2864
Consonant	reject same distribution	1	1	1	0
	significance level	2.4E-14	0	8.16E-09	0.1554
Tone	reject same distribution	1	1	1	1
	significance level	0.0152	0.0210	0.0138	0.0163

**Table 6.** Statistics of the words formed by different phonemes

		Monosyllable	Bisyllable	Trisyllable	Quadrisyllable	Average
Vowel	SD/mean	0.5924	0.6523	0.6765	0.6504	0.6429
	Max/Min	7.48	7.2972	7.0037	7.9027	7.4209
Consonant	SD/mean	0.3528	0.5118	0.5743	0.6406	0.5199
	Max/Min	4.4	6.6366	7.7454	10.1056	7.2219
Tone	SD/mean	0.1302	0.2327	0.2863	0.2421	0.2228
	Max/Min	1.3740	1.7626	2.0495	1.8616	1.7619

Chinese. The high degree of sparsity of words also leads to a high degree of sparsity of sentences, which are strings of words. The high degree of sparsity not only leaves enough room for new words to appear but contributes to effective and reliable communication: if all theoretically possible combinations were used in a language, one single mistake in communication would lead to misinterpretation, which would render error correction virtually impossible and would lower the efficiency of communication.

### 3.4. The Symmetry of DC

Formula (1) implies that the DCs of two phonemes in a contrast are not necessarily the same. Hence we conducted t tests on all vowels, consonants and tones in contrast to see whether two contrastive phonemes have significantly different DCs. The results are given in table 5. They show that the two phonemes in most contrasts have significantly different DCs. We statistically examined the number of words those phonemes occur in. Owing to the small sample size (there are only 6 vowels, 5 tones and 22 consonants), the statistical test on the distribution of the number of words constituted by these phonemes can hardly yield reliable results. Therefore, we calculated the coefficient of variance of the number of words (coefficient of variation = standard deviation/mean) and the ratio of the maximum to the minimum. The results are shown in table 6.

As can be seen in table 6, for both vowels and consonants in contrasts, there is a significant difference between the numbers of words in which two contrastive phonemes appear. The maximum can be seven times as large as the minimum! In contrast, the tones show much smaller disparity: the maximum is just about 1.7 times as much as the minimum.

#### 4. Discussion and Conclusion

Our study has obtained the following results:

(a) The DC of phonemes reflects the relation between phonemes. The DC of Mandarin Chinese consonants, vowels and tones declines exponentially with the increase in the number of syllables in a word, rapidly approaching zero.

(b) Compared with other consonants, the contrastive consonants that differ in only one distinctive feature and the consonants that are absent in some Chinese dialects have significantly higher DC.

(c) In Mandarin Chinese, there are no significant differences in DC between the contrastive consonants differing in one distinctive feature and the consonants absent in some Chinese dialects.

(d) Mandarin Chinese phonemic combinations exhibit a high degree of sparsity, which grows exponentially with the increase of the number of syllables in a word and rapidly approaches 1.

(e) With regard to DC, there is asymmetry between the two phonemes in a contrast. Therefore, the study of the influence of the contrast on the two phonemes requires that the two phonemes should be separately assessed.

A high degree of sparsity and a low DC render it possible to correctly recognize one particular word with only a fraction of its phonemes, which implies that phonemic mistakes in a speaker's utterance will not necessarily ruin communication. In other words, a high degree of sparsity reduces the demand on correct articulation and sound discrimination, not only relieving the psychological burden during linguistic communication, but providing redundant cues for a language system.

Sounds are in constant change during language evolution: some phonemic contrasts disappear when one phoneme in the contrast disappears. In some Chinese dialects there are no such consonants as /l/, /tʂ/, /tʂʰ/, /ʂ/ and /x/, which means that there are no contrasts such as /n/-/l/, /ts/-/tʂ/, /tsʰ/-/tʂʰ/, /s/-/ʂ/ and /f/-/x/. However, Surendran and Niyogi [2006] pointed out that there is no proof that their role (i.e. the functional load) is less important than that of other consonants. Our study shows that in Mandarin Chinese, the DCs of these consonants are roughly the same as that of contrastive consonants differing from each other in one distinctive feature, and higher than that of the other consonants. What is more, these consonants occur in no fewer words than other consonants. All these results point to one fact: the role and importance of phonemic contrasts are not the main cause for their disappearance. Physical, physiological, social and geographic factors may also lead to phonemic changes and the loss of some phonemic contrasts [Baranowski, 2010; Dinkin, 2009; Labov, 1994, 2001; Ohala, 1981, 1997].

The contrastive phonemes differing in only one distinctive feature show significantly higher DC than other contrastive phonemes, a discovery indicating that in a language system, the contrast between distinctive features may be more important, for which we have not yet found a satisfactory explanation.

Based on distinctive feature detection, automatic speech recognition systems identify, via inspecting acoustic events and acoustic landmarks, the rough phonetic classes and distinctive features of some segments, which serve as the basis for speech recognition. In these processes, some distinctive features may be misidentified, for the corresponding algorithms identify distinctive features on the basis of some acoustic parameters, which are liable to contextual influences.

Our study shows that, because of the redundant cues, one phonemic mistake is unlikely to lead to misrecognition of a word, especially when the word has more than one syllable. Therefore, automatic speech recognition based on the detection of phonemic features may, at least theoretically, achieve better results via recognizing a word as a whole, namely, recognizing the whole sequence of phonemic distinctive features of this word, than through a hierarchy of features, phonemes, syllables and words [Jansen and Niyogi, 2010; Stevens, 2002].

Our investigation into a number of statistical parameters of Mandarin Chinese phoneme networks reveals that the speech sound system tends to be reliable and energy saving, which renders it an efficient and effective communication system [Yu et al., 2011]. In this paper, by investigating Chinese phonemic contrast, we found that Chinese phonemic combinations have high degrees of sparsity and low DC, which can both contribute to effective and reliable communication and reduce the demand on speech production as well as perception. Once again, the balance between efficiency and reliability in a language system is demonstrated.

## Acknowledgments

We are grateful to Dr. Yi Xu and two anonymous reviewers for their detailed comments and suggestions. This work is partly supported by the National Social Science Foundation of China (No. 09BYY024 and 11&ZD188).

## References

- Atkinson, Q.D.: Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science* 332: 346–349 (2011).
- Baranowski, M.: On the role of social factors in the loss of phonemic distinctions. *Univ. Pa. Working Papers Ling.* 16: article 2 (2010).
- Carter, D.M.: An information-theoretical analysis of phonetic dictionary access. *Computer Speech Lang.* 2: 1–11 (1987).
- Esser, K.H.; Condon, C.J.; Suga, N.; Kanwal, J.S.: Syntax processing by auditory cortical neurons in the FM-FM area of the mustached bat *Pteronotus parnellii*. *Proc. Natl. Acad. Sci. USA* 94: 14019–14024 (1997).
- Dinkin, A.J.: *Dialect Boundaries and Phonological Change in Upstate New York*; diss. University of Pennsylvania, Philadelphia (2009).
- Duanmu, S.: *Syllable structure: the limits of variation* (Oxford University Press, Oxford 2009).
- Hay, J.; Bauer, L.: Phoneme inventory size and population size. *Language* 83: 388–400 (2007).
- Hockett, C.: The origin of speech. *Sci. Am.* 203: 88–111 (1960).
- Hockett, C.F.: *A manual of phonology* (Indiana University Press, Bloomington 1955).
- Hoeningwald, H.M.: *Language change and linguistic reconstruction* (University of Chicago Press, Chicago 1960).
- Ingram, D.: *First language acquisition: method, description and explanation* (Cambridge University Press, Cambridge 1989).
- Jackendoff, R.: *Foundations of language* (Oxford University Press, Oxford 2002).
- Jansen, A.; Niyogi, P.: Detection-based speech recognition with sparse point process models. *Proc. ICASSP*, pp. 4362–4365 (2010).
- King, R.D.: Functional load and sound change. *Language* 13: 831–852 (1967).
- Labov, W.: *Principles of linguistic change: internal factors* (Blackwell, Oxford 1994).
- Labov, W.: *Principles of linguistic change: social factors* (Blackwell, Oxford 2001).

- Manica, A.; Amos, W.; Balloux, F.; Hanihara, T.: The effect of ancient population bottlenecks on human phenotypic variation. *Nature* 448: 346–348 (2007).
- Martinet, A.: Remarques sur le système phonologique du français. *Bull. Soc. Ling. Paris* 34: 191–202 (1933).
- Martinet, A.: Économie des changements phonétiques (Francke, Bern 1955).
- Mathesius, V.: La structure phonologique du lexique du tchèque moderne. *Trav. Cercle ling. Prague* 1: 67–84 (1929).
- Meyerstein, R.S.: Functional load: descriptive limitations, alternatives of assessment and extensions of application (Mouton, The Hague 1970).
- Ohala, J. J.: Articulatory constraints on the cognitive representation of speech; in Myers, Laver, Anderson, The cognitive representation of speech, pp. 111–122 (North Holland, Amsterdam 1981).
- Ohala, J.J.: Aerodynamics of phonology. *Proc. 4th Seoul Int. Conf. on Ling.*, pp. 92–97 (1997).
- Peeters, B.: Diachronie, phonologie et linguistique fonctionnelle (Peeters, Louvain-la-Neuve 1992).
- Stevens, K.N.: Acoustic phonetics (MIT Press, Cambridge 1998).
- Stevens, K.N.: Toward a model for lexical access based on acoustic landmarks and distinctive features. *J. acoust. Soc. Am.* 111: 1872–1891 (2002).
- Surendran, D.; Niyogi, P.: Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals; in Nedergaard Thomsen, Competing models of language change: evolution and beyond, pp. 43–58 (Benjamins, Amsterdam 2006).
- Teich, M.C.; Lowen, S.B.: Fractal patterns in auditory nerve-spike trains. *IEEE Eng Med Biol Mag* 13: 197–202 (1994).
- Trubetzkoy, N.: Grundzüge der Phonologie. *Trav. Cercle ling. Prague* 7 (1939).
- Yu, S.; Liu, H.; Xu, C.: Statistical properties of Chinese phonemic networks. *Physica A* 390: 1370–1380 (2011).
- Zuidema, W.; de Boer, B.: The evolution of combinatorial phonology. *J. Phonet.* 37: 125–140 (2009).