Physica A 390 (2011) 1370–1380

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Statistical properties of Chinese phonemic networks

Shuiyuan Yu [a], Haitao Liu [b,*], Chunshan Xu [c]

[a] School of Computer, Communication University of China, Beijing, 100024, China
[b] School of International Studies, Zhejiang University, Hangzhou, 310058, China
[c] Institute of Applied Linguistics, Communication University of China, 100024 Beijing, China

ARTICLE INFO

Article history:
Received 6 September 2010
Received in revised form 21 October 2010
Available online 22 December 2010

Keywords:
Phonemic network
Complex network
Speech sound system

ABSTRACT

The study of properties of speech sound systems is of great significance in understanding the human cognitive mechanism and the working principles of speech sound systems. Some properties of speech sound systems, such as the listener-oriented feature and the talker-oriented feature, have been unveiled with the statistical study of phonemes in human languages and the research of the interrelations between human articulatory gestures and the corresponding acoustic parameters. With all the phonemes of speech sound systems treated as a coherent whole, our research, which focuses on the dynamic properties of speech sound systems in operation, investigates some statistical parameters of Chinese phoneme networks based on real text and dictionaries. The findings are as follows: phonemic networks have high connectivity degrees and short average distances; the degrees obey normal distribution and the weighted degrees obey power law distribution; vowels enjoy higher priority than consonants in the actual operation of speech sound systems; the phonemic networks have high robustness against targeted attacks and random errors. In addition, for investigating the structural properties of a speech sound system, a statistical study of dictionaries is conducted, which shows the higher frequency of shorter words and syllables and the tendency that the longer a word is, the shorter the syllables composing it are. From these structural properties and dynamic properties one can derive the following conclusion: the static structure of a speech sound system tends to promote communication efficiency and save articulation effort while the dynamic operation of this system gives preference to reliable transmission and easy recognition. In short, a speech sound system is an effective, efficient and reliable communication system optimized in many aspects.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

A speech sound system is an essential human communication system. The study of its properties can throw light on its operation, on its relations with other cognitive abilities of human mind and on the nature of human languages. The exploration into speech production and recognition and the study of the linguistic universals [1–4] found in various languages have uncovered two linguistic properties: the talker-oriented feature, which refers to the fact that the phonemes of a speech sound system always tend to effect successful communication with the least articulation effort, and the listener-oriented feature, which refers to the requirement that the phonemes of a speech sound system be easy to perceive and discriminate. There are many speech theories addressing speech properties, including Distinctive Feature Theory [5], Quantal Theory [6], Dispersion Theory [7,8] and Auditory Enhancement Hypothesis [9,10], etc.

* Corresponding author.
E-mail addresses: shuiyuan.y@gmail.com (S. Yu), lhtzju@gmail.com (H. Liu), adinxu@yahoo.com.cn (C. Xu).

0378-4371/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2010.12.019

The speech sound system consists of rules that prescribe allowable phonemes and phoneme combinations. These phonemes organize into a consistent whole. Hence, it is better to study the phoneme properties from the perspective of this whole [11]. At the same time, the speech sound system is a communication system functioning to choose phonemes and phoneme combinations to communicate information. Therefore, the working properties of a speech sound system are also important. This paper, treating the Chinese phonemes as a coherent whole, explores into the speech sound system through a comparison between its static rules and its properties exhibited in dynamic operation.

A language system is a network with a complex structure [12]. Many researches, which all adopt complex network approaches, have examined properties of languages at various levels, including the study of phonetic and lexical network [13–18], the study of syntactic network [19–25] and the study of semantic network [26–29]. These studies have discovered many important linguistic properties. However, the study of speech sound systems by means of phonemic network is so far rather rare—another motivation of the research presented in this paper, which we hope may contribute to the understanding of human speech sound systems.

Based on the six Chinese full phonemic networks (each comprises the entire list of Chinese phonemes) constructed in our research, we have measured, of the networks extracted from real corpora, the weighted parameters and the robustness against attacks and errors to study the dynamic properties of the speech sound system. Moreover, to examine the static properties of the speech sound system, we have also conducted a statistical study of the dictionaries concerned, and measured the degrees of the networks extracted from real corpora and the parameters of networks based on above-mentioned dictionaries. The following section is to introduce the construction of Chinese phonemic networks in our research. The third section will present the measuring and the statistical approaches employed in the research. The last section will report the results and the discussion.

2. Structure of phonemic network

A Chinese speech sound system is a hierarchy, comprising phonemes, Chinese characters (monosyllables), words (polysyllables) and sentences (polysyllables). In other words, in terms of speech, Chinese sentences comprise words, which consist of characters, which, as monosyllables, are ultimately composed of phonemes. Since Chinese characters are meaningful—in fact, many characters can stand alone as words (monosyllable words), we respectively constructed the full phonemic networks of Chinese characters (monosyllables), words (polysyllables) and sentences in order to, on the one hand, investigate the properties of speech sound systems, and on the other hand, compare the properties of speech sound systems at different linguistic levels.

Corpus: Our research has employed four materials to construct phonemic networks: the dictionary of 7290 fundamental Chinese characters (monosyllables), the dictionary of 88 250 commonly used words (roughly 163 500 characters), tourism news (roughly 40 000 characters, 3650 clauses and 6240 different words) and daily dialogues (roughly 6000 characters, 840 clauses and 1200 different words) extracted from Chinese conversation textbooks available at http://www.china.org.cn/learning_chinese/dialogue/.

Node: the nodes of networks are Chinese phonemes, which total 46, including 21 vowels, 2 semi-vowels and 23 consonants.

Edge: the edges in the networks are all directed ones linking neighboring phonemes. For example, Chinese character '甘' (/kan/ sweet) is composed of 3 phonemes /k/, /a/ and /n/, which are linked by 2 directed edges: /k/-/a/ and /a/-/n/.

Different choices of granularity levels in segmenting linguistic materials will produce networks at different linguistic levels and bring to these networks different amounts of edges and different node degrees, resulting in networks with different properties. We segment the linguistic materials of real text at two levels of granularity: the level of sentence and the level of word. Thus, plus the networks based on the two dictionaries, we have six phonemic networks. Fig. 1 is a sample of phonemic networks. The graph (a) is the phonemic network of the sentence '他给我打电话' (he calls me on the telephone), (b) is the phonemic network of the words extracted from the same sentence after segmentation, and (c) is the phonemic network of the monosyllables (characters) that constitutes the sentence. For the sake of a convenient computerized editing, we adopt SAMPA (Speech Assessment Methods Phonetic Alphabet, http://www.elgin.free-online.co.uk/), instead of the notations of IPA. It can be seen in this figure that the differences between (a) and (b), (b) and (c) networks lie in the fact that in the network at the level of word, neighboring boundary phonemes of adjacent words in a sentence, such as '他' (he), '给' (to), '我' (me), '打' (call) and '电话' (telephone), don't link together, however, in the network at the level of sentence, they do; and in the network at the level of character, the neighboring boundary phonemes of adjacent characters in a word, such as '电', '话', don't link together, but, in the network at the level of word, they do. In this figure, every edge in three graphs is labeled with a weight.

In this paper, these six networks are respectively named: PNW (phonemic network of words), PNUS (phonemic network of monosyllables (characters)), PNDS (phonemic network of dialogue sentences), PNDW (phonemic network of dialogue words), PNNS (phonemic network of news sentences) and PNNW (phonemic network of news words). In these six phonemic networks, PNUS is the network of monosyllables, and the others are the networks of polysyllable.

To study the actual use of Chinese, we have selected two texts of different styles to investigate the influence of styles on the speech sound system. The following section will present the measuring and the statistical approaches employed in the research.

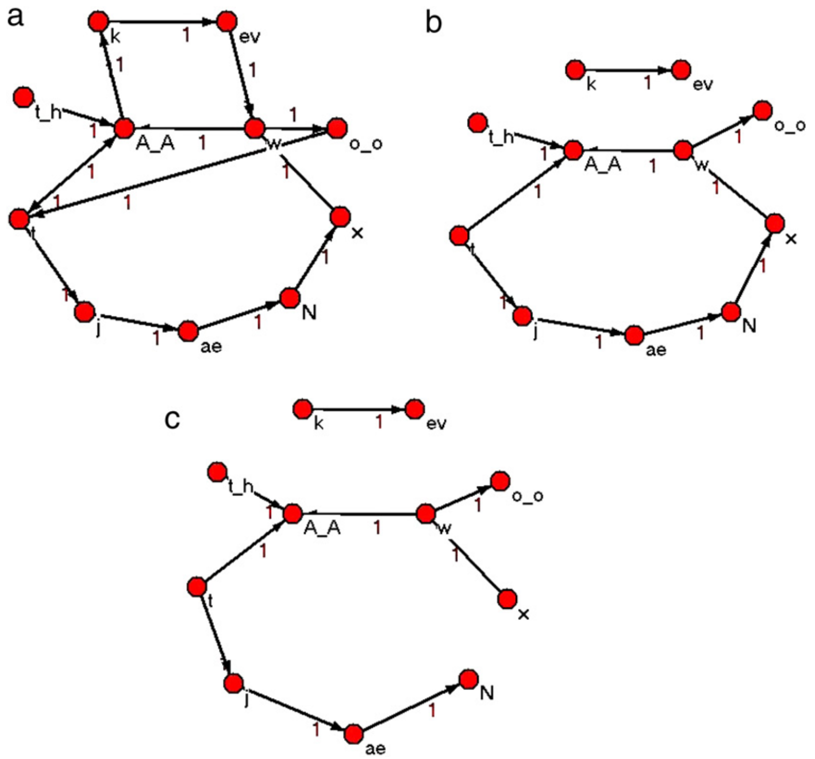

Fig. 1. A sample of phonemic networks. (a) is the phonemic network of the sentence '他给我打电话', (b) is the phonemic network of the words extracted from this sentence after segmentation, and (c) is the phonemic network of the monosyllables (characters) that constitutes the sentence.

Table 1
Basic statistics for six phonemic networks and two random networks.

Network	Mean vertex–vertex distance L	Mean degree $\langle k \rangle$	Connectivity	Total number of edges	Clustering coefficient C
PNW	1.88	26.39	1.00	611	0.19
PNUS	1.54	8.74	0.19	201	0.03
PNDS	1.97	22.61	1.00	520	0.11
PNDW	2.10	19.17	1.00	441	0.07
PNNS	1.89	25.74	1.00	592	0.19
PNNW	1.93	24.13	1.00	555	0.14
ER model	1.76	24.04	1.00	553	0.27
SF model	1.45	24.43	0.47	562	–

3. Measurement of phenomic network

Measuring statistical parameters of networks is the fundamental work in network research. The comparison among the parameters of different networks can reveal the similarities and differences among networks and thus make for a better understanding of them.

3.1. Statistical properties of networks

Table 1 lists out major statistical parameters of the six phonemic networks and two random networks with similar size. There are two major types of complex networks: exponential networks and power law networks. For the sake of a comparison between the phonemic networks and the typical complex networks, Table 1 gives the statistical parameters of a random ER network [30] and a random SF network [31], which both have the same amount of nodes as and similar average degree to the phonemic networks.

Newman [32] has reviewed the history of complex network research, presenting, in his paper, major statistical parameters of 27 complex networks in such fields as society, information, technology and biology, etc. A comparison of Table 2 in Newman's work and Table 1 in this paper shows that the average degree of phonemic networks reported in this paper is very high in these fields. Besides, the average distance of phonemic networks is the shortest. This implies a very strong combining capacity in nodes of phonemic networks. In other words, the same amounts of links in phonemic networks involve fewer nodes than in other networks, which is an indication of high efficiency of phonemic networks in exploiting nodes.

Table 2
Six networks' significance probabilities of normal and exponential distribution tests.

Network	Normal	Exponential
PNW	.189	
PNUS		.073
PNDS	.398	
PNDW	.503	
PNNS	.347	
PNNW	.515	

Table 3
The coefficients of determination of power law fitting.

Network	PNW	PNUS	PNDS	PNDW	PNNS	PNNW	Average
Adj-R^2	0.93	0.89	0.91	0.91	0.78	0.91	0.89
Index	−1.672	−1.433	−1.741	−1.649	−1.167	−1.681	−1.557

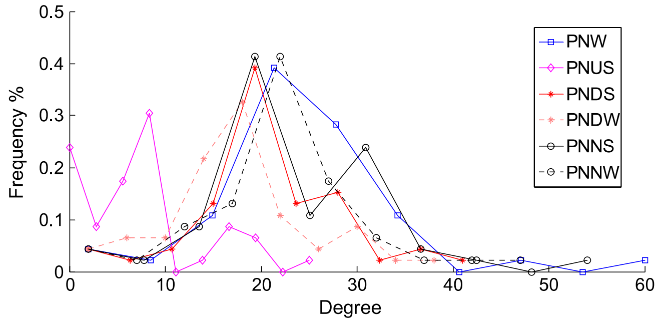

Fig. 2. Distributions of degrees of six networks.

Degree distribution, in addition to average degree, is another important property of networks. What characteristics then can be found in the degree distribution of phonemic networks?

3.2. The distribution of node degree

In phonemic networks, node degree, which means the amount of nodes to which one node directly links, signifies the importance of phoneme nodes in phonemic networks; and weighted degree, which refers to the frequency that a node gets linked when networks are in operation, indicates the importance of a node in the operation of a network. In a nutshell, the distribution of weighted degree demonstrates the priority enjoyed by some phonemes in operation while, determined by speech rules, degree distribution displays the structure of a phonemic network and allowable links between phonemes

We divided the existential interval of degrees evenly into ten sub-intervals and measured the proportion of the nodes falling in each sub-interval. The measurement results are shown in Fig. 2.

Since each curve of five networks of polysyllables in Fig. 2 has a single peak, we conducted normal distribution test on them with 95% confidence level. The curve of the degree of PNUS shows that these nodes with less degree are inclined to have more frequency, so we conducted exponential distribution test on PNUS with 95% confidence level. Table 2 gives the significance probabilities greater than 0.05 (an indication of an acceptable hypothesis). Apparently, each degree distribution besides PNUS stands normal distribution test, which means that, in the five phonemic networks of polysyllables, the peak value of degree distribution is around the mean. As the node degree increases, the amount of corresponding nodes decreases dramatically—a sign of the fact that the nodes in the five networks are homogeneous because they possess roughly the same amount of links. The PNUS stands exponential distribution test, it means that PNUS is not homogeneous.

The distribution of weighted degree is measured with the same approach adopted in measuring the distribution of degree. Owing to the fact that these nodes with less weighted degrees present themselves more frequently, we fit their accumulative probability distribution functions to power law function. The coefficients of determination are shown in Table 3. In contrast with the balanced unimodal distribution of degree, the distributions of weighted degree in all six networks accord with power law distribution. The power law fitting index of six networks are rather similar, with a mean of −1.557 and a variation coefficient (the indices of deviation from mean) of −0.14. This is evidence that the six networks have similar weighted degree distributions. A further comparison among these networks will be pursued in following sections.

As to the networks based on real text, the degree of a node indicates the combining capacity of one phoneme while its weighted degree signifies the frequency of its employment in actual use. The unimodal and balanced degree distribution approximates normal distribution—an indication of the fact that in these networks the linking capacities of the nodes, which

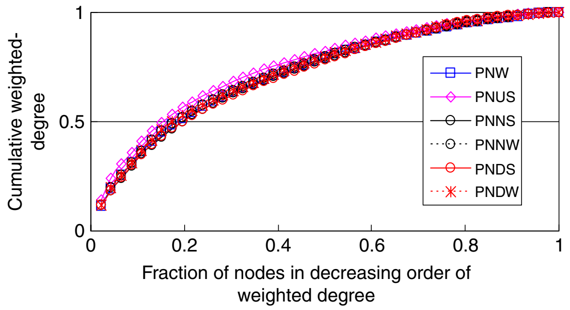

Fig. 3. An accumulative probability of weighted degrees in decreasing order.

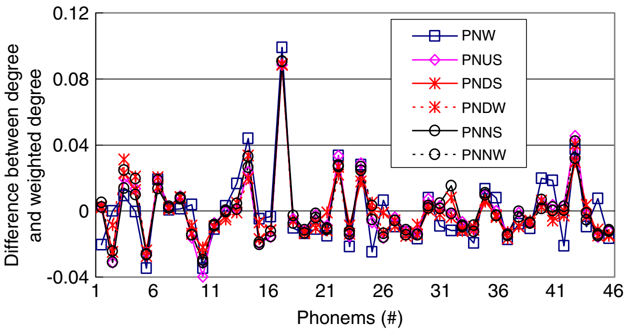

Fig. 4. The difference between degree and weighted degree of phonemes.

Table 4
The Average normalized weighted degree of vowels and consonants of five polysyllabic networks.

Networks	Size of text (character)	Granularity level	Average weighted degree of vowels	Average weighted degree of consonants
PNDW	6 000	Word	2.07	1.59
PNNW	40 000	Word	2.06	1.62
PNW	163 500	Word	2.02	1.67
PNDS	6 000	Sentence	2.04	1.69
PNNS	40 000	Sentence	2.03	1.71
Mean			2.04	1.65

are somewhat homogeneous, all cluttering around the average degree. At the same time, the weighted degree obeys power law distribution, namely, a small portion of phonemes are widely used while numerous others are rarely. Fig. 3 presents the graph of accumulative probability of network nodes in terms of weighted degree. In this graph, x-axis is the fraction of nodes in decreasing order of weighted degree and y-axis is the accumulative probability. We can see that the accumulative probability of less than 20% nodes exceeds 50% while all the rest 80% nodes only account for an accumulative probability of less than 50%. However, in terms of degree, the accumulative probability of 50% covers nearly 40% nodes, which shows that node degrees are rather evenly distributed. In short, phonemes enjoy different priorities in actual use.

Both Table 3 and Fig. 3 reveal that a small portion of phonemes enjoys priority in the actual working of speech sound systems. As an effort to identify those phonemes, we measure, among phonemes, the difference in degree and weighted degree. The result is shown in Fig. 4, in which, for the sake of a convenient comparison, degrees and weighted degrees are normalized in following ways: normalized degree = degree/totality of degrees of all phonemes; normalized weighted degree = weighted degree/totality of weighted degrees of all phonemes. As shown in Fig. 4, the six networks have similar curves and so exhibit similar regularities. In order to investigate the phonemes that take priority over the others in the actual operation of speech sound systems, we compute the average normalized weighted degree of vowels and consonants in five polysyllabic networks, the results are shown in Table 4. The average normalized weighted degrees of vowels in the five networks are similar, the granularity level in segmenting linguistic materials and the size of linguistic materials (or sample) do not seriously influence them. These factors do not also influence the average normalized weighted degrees of consonants. T-test shows that the probability value is approximately 0.000, less than 0.05. The 95% confidence interval of the difference is 0.3324–0.4436, without including zero. Therefore, the two means are significantly different. In other words, the frequency of vowels is promoted in actual use, or, the vowels take priority over consonants in actual use.

In a continuous speech flow, vowels carry more acoustic energy than consonants, a property that, together with the harmonic structure of vowels, makes for easy discrimination and good transmission. Consequently, the priority of vowels suggests good transmission and easy recognition. Nevertheless, since the production of sounds of higher energy physically

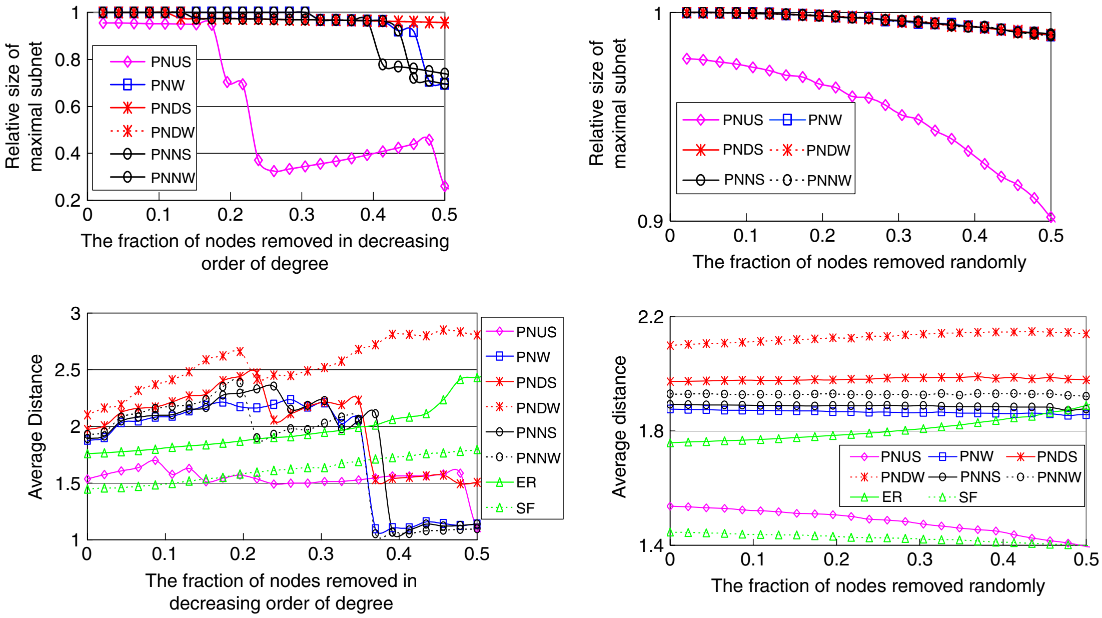

Fig. 5. Changes in the average distance and relative size of maximal subnet of the network as a function of the fraction of the removed nodes randomly and in decreasing order of degree.

calls for more effort, the priority of vowels obviously goes against the least effort principle if, in terms of the strength of human beings, the energy consumption difference between vowel and consonant articulation is significant.

So far, our concern is with the statistical parameters of networks in normal operation. The following section will be committed to the research of robustness of networks against attacks and errors. We first delete some nodes in the networks, then observe the responses of networks and analyze the interrelations between the responses and the structure of speech sound systems to find out the anti-interference properties of speech sound systems.

3.3. The robustness of phonemic networks

Presently, most anti-interference researches concerning speech transmission probe, from acoustic or auditory perspective, into only a limited group of sounds. Our research concerns the influence of malfunction of one or more phonemes on the entire phonemic network, namely, the changes in network properties resulting from the elimination of some nodes. These changes have close relations with network structures.

By measuring network diameters and relative sizes of maximum sub-networks, Albert examined the effects of targeted attacks and random errors on complex networks, whose node degree usually has two typical distributions: exponential distribution and power law distribution [33]. They found that, owing to homogeneity of its nodes, targeted attacks and random errors bring about similar consequences in an exponential network while in a power law network, whose nodes are heterogeneous, the targeted attacks lead to much more serious consequences than random errors.

The network diameter, which can be defined as the average length of the shortest paths between any two nodes in the phonemic network, indicates the direct collocating relation among phonemes. A greater diameter demands more phonemes in successful communications, which naturally results in lower efficiency. The relative size of the largest connected components reflects the ability of a network to maintain operating under attacks. These two parameters are the measures of overall communicating capacity of a network. Targeted attacks can be considered as articulation difficulties and random errors as the random noise interferences in the course of speech transmission. Fig. 5 presents the variation curves of network diameters and relative sizes of the largest connected components when the networks suffer from targeted attacks and random errors.

In this figure, Δ with solid line is the sign for ER random network and Δ with dot line for SF random network. For any given number of nodes removed from the network, we have repeated the removing operation 3000 times to obtain an average reported in this figure. It is obvious that, as the result of similar degrees of nodes, the removal by decreasing order of degree and random removal do not bring about much different consequence, which conforms to the finding of Albert and his colleagues. Moreover, the high degree in phonemic networks also leads to, within a considerable range (roughly 40% for five polysyllable networks and around 18% for monosyllable network), virtual fixedness of the relative size of the largest connected components. However, though network diameter virtually does not change with random errors, it does increase dramatically under targeted attacks. This shows that articulation difficulties, though having no influence on the communicative capacity, do deteriorate the efficiency of phonemic networks while random noise interference actually has no influence at all on both communicative capacity and efficiency of phonemic networks.

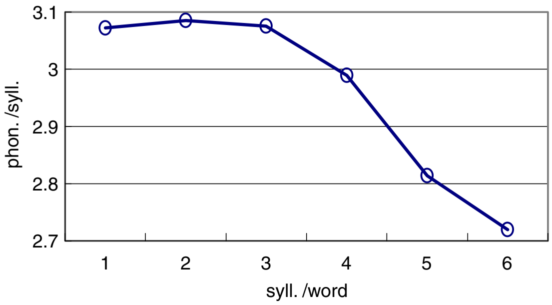

Fig. 6. The relation between word length and syllable length as found in the dictionary.

To make certain whether the anti-interference capacity is affected by the priority of some phonemes in the operation of speech sound systems, we have measured network diameters and the relative sizes of the largest connected components when nodes are removed by decreasing order of weighted degree. The results register that the variation of these two parameters lies between removal by decreasing order of degree and random removal. It means that the priority of some phonemes in the speech sound system operation does not influence the anti-interference capacity of speech sound systems.

In Fig. 5, the variations of the parameters in the two random networks follow the same pattern, which suggests that this attack-resistant capacity is not due to network structure but the high node degrees.

Languages are hierarchically structured, with the sub units composing the super ones. The relations between super and sub units may reflect the properties of linguistic system. Therefore, we have conducted researches into the compositional relations between Chinese words and syllables.

3.4. The relation between lexical complexity and syllabic complexity

The length of a word can be measured with its syllables and the length of a syllable is measured with its phonemes. Fig. 6 presents the relations between word length and syllable length as found in the dictionary of 88 250 commonly used words. In the figure, the maximum word length is prescribed to be six syllables because Chinese word length rarely exceeds six characters. Sometimes names for institution or the transcriptions of foreign geographical terms are rather long, but strictly speaking, they are more phrases than words.

Fig. 6 shows that the increase in word length is accompanied by a monotonous decrease in the average length of its syllables. That is, the longer a word is, the shorter the compositional syllables of this word are, or, the less phonemes each syllable averagely has. This finding is in accordance with Menzerath Law [34,35].

Fenk et al. [36] has conducted similar measurements in Italian, German and Indonesian. The results of Italian are rather similar in detail to Fig. 6: both have a maximum of about 2. It is advisable to fit it to cubic polynomial. The coefficient of determination is 0.961 and the peak value appears when the word length is 1.94 syllables. The normal fitting is also applicable to our data. The coefficient of determination to normal curve is 0.961 and the peak value appears when the word length is 1.79 syllables.

Fig. 6 shows that the more syllables a words has, the less phonemes each syllable averagely has, which signifies the tendency in a speech sound system to express longer words with less phonemes—a property contributing to high communication efficiency.

Apart from word length, the optimization of distributions of unit length at various linguistic levels may also promote communication efficiency. This is the motivation of our research of the distributions of word length and syllable length.

3.5. The distributions of word length and syllable length

Based on the dictionary of 88 250 commonly used words, we have studied the distribution of syllable length, and the results, presented in Fig. 7, show the fact that 2-syllable words account for the biggest share of Chinese words (around 54.12%) and the tendency that longer words account for smaller portion of Chinese words. The 403 toneless syllables in Chinese, when qualified with the four tones, provide 1612 available toned syllables in the Chinese speech sound system. Therefore, every toned syllable, in average, is shared by 3.49 monosyllable word: there are 5620 monosyllable words in Chinese. In other words, the ratio of monosyllable words to toned-syllables is 3.49($5620/1612 = 3.49$), much higher than the 0.0184 ratio of bi-syllable words to bi-syllables ($47\,812/(1612 * 1612) = 0.0184$). It is safe, therefore, to claim that there exists a tendency in Chinese to give priority to short words.

Chinese words are composed of monosyllabic characters. There are 403 syllables in Chinese, of which five syllables are composed of only one phoneme, 54 from two phonemes, 244 from three phonemes, and 110 from 4 phonemes. Theoretically, the 46 phonemes in Chinese, with the compositional rules and allophones left out, can possibly produce 23 monophonemic syllables (all vowels), 529 bi-phoneme syllables ($23\text{(consonants)} * 23\text{(vowels)} = 529$), 11 638 tri-phoneme syllables ($23 * 23 * 22 = 11\,638$) and 244 398 quadric-phoneme syllables ($23 * 23 * 22 * 21 = 244\,398$). The proportions of various

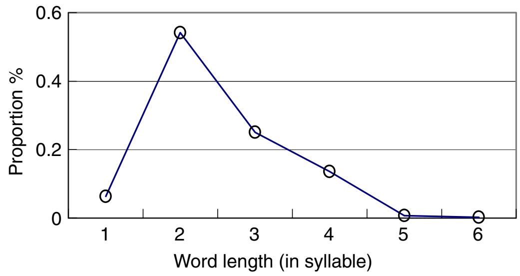

Fig. 7. The quantity distribution of word length.

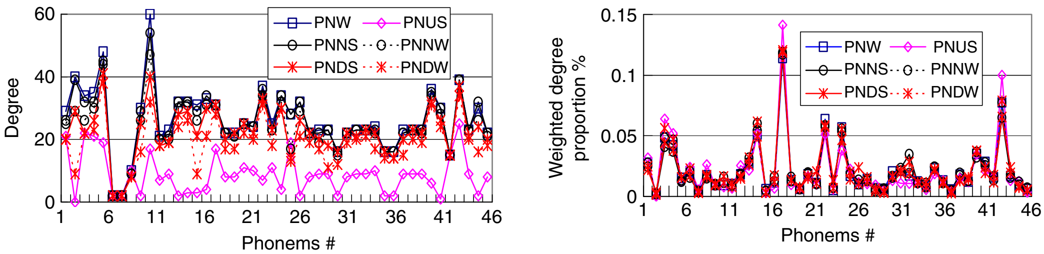

Fig. 8. Degree and weighted degree distributions of 46 phonemes in six networks.

syllables actually used in speech sound systems are respectively 0.2174(5/23 = 0.2174), 0.1021(54/529 = 0.1021), 0.0201(244/11 638 = 0.0201), and 0.0005(110/244 398 = 0.0005). Obviously, the same tendency as found in word length also exists in syllable length: monosyllables tend to comprise as less phonemes as possible, i.e. the Chinese speech sound system tends to use short syllables.

The priority of short words and short syllables, an advantage in terms of minimum articulation effort, reflects the optimization principle in the speech sound system regarding word formation.

In our study, we have used various materials to construct six phonemic networks concerning different linguistic levels. These networks exhibit similarities in many statistical properties. Then, are there topological similarities among them? To answer this question, we investigated the degree distributions of these six networks to detect the similarities and the differences because degree distribution is generally regarded as the indication of network structure.

3.6. Similarities among the six networks

The six networks constructed in our study are concerned with different linguistic levels. Nevertheless, all of them are polysyllable networks except for the network of fundamental characters, which is a monosyllable network. It has been learnt that these six networks share many similarities in degree distribution, weighted degree distribution and robustness against attacks and random errors, but it remains an unconcluded issue whether they are structurally similar. According to Fig. 8, we can see that, except for the network extracted from the dictionary of fundamental characters, all the networks have similar distribution curves. We conducted linear correlation fitting with regard to their degree and weighted degree. The result is shown in Table 5.

As shown in Table 5, the average coefficient of determination is 0.577, the maximum is 0.99 and the minimum is 0.05. The maximum coefficient of determination between the monosyllable network and other polysyllable networks is 0.16, which denies a linear correlation between them. Among the five polysyllable networks, the average coefficient of determination is 0.82, the maximum is 0.99 and the minimum is 0.57. The two coefficients less than 0.75 both derive from the networks extracted from dialogue corpus.

In Table 5, the average coefficient of determination of weighted degree in six networks is 0.994; the maximum and the minimum coefficients are respectively 0.99 and 0.86. This means all six networks, in actual operation, show similar preference for some phonemes though, owing to different styles, different corpora and different choices of granularity levels, the linking relations among nodes (phonemes) may differ in those networks.

In brief, all six phonemic networks exhibit similar statistical parameters. Nevertheless, their topologic structures may be different, especially the topological structure of monosyllable network, which displays tremendous difference from those of other networks. The possible reason may be that, though many Chinese characters are actually monosyllable words, the corpus based on dictionary of fundamental characters somewhat suggests a different linguistic level of analysis in comparison with other corpora. On the other hand, without exception, when these networks are in operation, only a small portion of nodes is frequently linked (used), a phenomenon pointing to a common working pattern shared by all six networks.

Table 5
The coefficients of determination of linear correlation fitting among six networks in terms of degree and weighted degree.

Coefficient of determination of degree	PNW	PNUS	PNDS	PNDW	PNNS	PNNW
PNW	1.00	0.10	0.84	0.57	0.99	0.92
PNUS		1.00	0.05	0.16	0.07	0.08
PNDS			1.00	0.78	0.87	0.96
PNDW				1.00	0.59	0.75
PNNS					1.00	0.93
PNNW						1.00
Coefficient of determination of weighted degree	**PNW**	**PNUS**	**PNDS**	**PNDW**	**PNNS**	**PNNW**
PNW	1.00	0.93	0.95	0.99	0.93	0.99
PNUS		1.00	0.86	0.94	0.87	0.95
PNDS			1.00	0.96	0.97	0.95
PNDW				1.00	0.94	0.99
PNNS					1.00	0.94
PNNW						1.00

4. Concluding remarks

We have investigated the statistical parameters of six phonemic networks, the compositional structure of words and syllables and the robustness of phonemic networks against targeted attacks and random errors. Our research has obtained the following results:

(a) In comparison with networks in other fields, phonemic networks have rather high node degrees and short average distances, which indicate high efficiency in exploiting nodes.
(b) The degrees of phonemic networks conform to exponential distribution while the weighted degrees obey power law distribution, and vowels' weighted degrees are higher than consonants'. These results show that in the actual operation of phonemic networks, nodes, though possessing roughly the same amount of static allowable links, are used (linked) with significantly different frequency—a conclusion supported by the evidence that vowels are used more frequently than consonant.
(c) Phonemic networks display high robustness against both targeted attacks (articulation difficulties) and random errors (random noise interferences). Random noise interferences have virtually no influence on both communication capacity and efficiency of networks and the articulation difficulties, though causing a loss in communication efficiency, neither do harm to communication capacity of networks.
(d) The statistical study of the dictionary of commonly used words registers the higher frequency of shorter words and syllables and the tendency that the longer a word is, the shorter the syllables composing it are.

All these results point to the tendency of speech sound systems to promote communication efficiency and facilitate sound articulation, transmission and recognition. In other words, a human speech sound system, which has been optimized in these aspects that are essential for any communication system, is quite successful.

Discussion:

Result (a) is the measurement of degrees of networks, which indicates the high efficiency of phonemic networks in exploiting phoneme nodes and reveals the structural properties of networks as well as the allowable links in a speech sound system. Result (d) comes from the statistical study of the dictionary of commonly used words, uncovering the higher frequency of shorter words and syllables and the tendency in word formation to use as few phonemes as possible; apart from that, it also signifies the homogeneity of the degrees of networks, which is the sign of the maximum entropy in phonemes employment. In short, all these properties and tendencies point to the fact that the phonemic rules of Chinese speech sound systems contribute to high communication efficiency.

Result (b) and result (c) are concerned with the operations of speech sound systems. Result (b) points out the priority that the speech sound system gives to vowels in operation, a property that, together with the harmonic structure of vowels, facilitates speech recognition and enhance the robustness against noise interference. Result (c) reveals the high robustness of phonemic networks against serious targeted attacks and random errors. In short, the operating properties and patterns of speech sound systems make for good transmission and easy recognition.

Briefly, when it comes to formation of words and syllables, speech sound systems put emphasis on the efficiency, while in the actual operation it attaches importance to the reliability of transmission and recognition.

In spite of the universal existence of Least Effort Principle in human behaviors [37,38], we have not observed, in the operation of speech sound systems, a direct evidence in favor of this principle except for the fact that high efficiency reduces the amount of phonemes needed. Before dwelling on the energy-saving properties of speech sound systems, we will first address the conditions that restrain the reduction of articulation effort in human speech communication systems.

The study of the continuous human speech flows shows that vowels generally last longer and cost more energy than consonants, which, according to the law of conservation of energy, proves that the articulation of consonants is more energy saving. In other words, to save energy means to use as few phonemes as possible and, when this requirement is met, to use as many consonants as possible.

As seen in preceding sections, the static rules of speech sound systems incline it to use as few phonemes as possible, a property contributing to energy saving in articulation. The dynamic operation of this system, nevertheless, presents a definite favor for vowels, which enhances the reliability of communication but goes against energy saving in articulation. To put it in another way, as far as Chinese is concerned, it is in choosing allowable phonemes and combinational rules that the Least Effort Principle, which is conducive of the reduction of time spent in articulation, plays its role. In actual operation, this principle will give way to compromise between energy saving and reliability of communication. In other words, since the Least Effort Principle has retreated from the scene at this stage, the speech sound system, with its attention shifted from energy saving to reliability, gives no priority to consonants.

The linking capacity of phonemes is rather strong: one phoneme may link many others. It is hence not likely to predict, when one phoneme occurs, its succeeding ones. In other words, the adjacent phonemes in a sound flow are comparatively independent. This property, as the price for high efficiency, reduces the redundancy of the system, impairs its anti-interference capacity and finally lowers its reliability. The solution to this problem provided by human speech sound systems is increasing system redundancy via co-articulation. Owing to co-articulation, there is no clear boundary between adjacent phonemes in a sound flow because a phoneme always merges with both its predecessor and its successor to form a continuous stream. In this way, the actual lasting time of one phoneme is lengthened, with part of its message merged into both its predecessor and its successor. Obviously, this solution increases temporal redundancy, enhances anti-interference capacity in transmission and facilitates sound recognition.

Easy sound recognition is consistent with high reliability in transmission because (1) it is easy to discriminate two sounds and unlikely to confuse them in transmission if they exhibit sufficient acoustic difference, and (2) since vowels carry relatively high energy and their frequency at the formant is low, they can better resist attenuation and noise interference in transmission. In brief, what makes for high reliability also contributes to sound recognition—a phenomenon resulting from the evolutionary adaptation of human auditory system to environment.

In conclusion, the allowable phonemes and compositional rules reflect high efficiency in exploiting phonemes, which also contributes energy saving and easy recognition while the dynamic networks in actual operation present the properties of high reliability and easy recognition. In other words, high reliability, energy saving and easy recognition all make for high efficiency of speech sound systems. Therefore, the speech sound system is an efficient and effective communication system that has achieved balanced optimization in the course of evolution.

Acknowledgements

We thank the referees for insightful comments. This work was supported by the National Social Science Foundation of China (09BYY024) and Communication University of China as one of "211" Key Projects.

References

[1] I. Maddieson, Patterns of Sound, Cambridge University Press, 1984.
[2] R. Carré, "Speaker" and "Speech" characteristics: a deductive approach, Phonetica 51 (1994) 7–16.
[3] R.L. Diehl, Acoustic and auditory phonetics: the adaptive design of speech sound systems, Phil. Trans. R. Soc. B 363 (2008) 965–978.
[4] B. Lindblom, P. MacNeilage, M. Studdert-Kennedy, Self organizing processes and the explanation of language universals, in: Brian Butterworth, Bernard Comrie, Östen Dahl (Eds.), Explanations for language universals, Walter de Gruyter & Co, 1984, pp. 181–203.
[5] R. Jakobson, G. Fant, M. Halle, Preliminaries to Speech Analysis, MIT Press, Cambridge, 1952.
[6] K.N. Stevens, The quantal nature of speech: evidence from articulatory–acoustic data, in: E.E. David, P.B. Denes (Eds.), In Human Communication: A Unified View, McGraw-Hill, New York, NY, 1972, pp. 51–66.
[7] J. Liljencrants, B. Lindblom, Numerical simulation of vowel quality systems: the role of perceptual contrast, Language 48 (1972) 839–862.
[8] B. Lindblom, Phonetic universals in vowel systems, in: J.J. Ohala, J.J. Jaeger (Eds.), Experimental phonology, Academic Press, Orlando, FL, 1986, pp. 13–44.
[9] R.L. Diehl, K.R. Kluender, On the objects of speech perception, Ecol. Psychol. 1 (1989) 121–144.
[10] R.L Diehl, K.R. Kluender, Reply to commentators, Ecol. Psychol. 1 (1989) 195–225.
[11] B. Lindblom, O. Engstrand, In what sense is speech quantal? J. Phonetics 17 (1989) 107–121.
[12] R.A. Hudson, Language Networks: The New Word Grammar, Oxford University Press, Oxford, 2007.
[13] A. Mukherjee, M. Choudhury, A. Basu, N. Ganguly, Self-organization of the sound inventories: analysis and synthesis of the occurrence and co-occurrence networks of consonants, J. Quantit. Linguist. 16 (2) (2009) 157–184.
[14] J. Li, J. Zhou, Chinese character structure analysis based on complex networks, Physica A 380 (2007) 629–638.
[15] Y. Li, et al., Small-world patterns in Chinese phrase networks, Chinese Sci. Bull. 50 (2005) 286–288.
[16] G. Peng, J.W. Minett, W.S.-Y. Wang, The networks of syllables and characters in Chinese, J. Quantit. Linguist. 15 (3) (2008) 243–255.
[17] S. Arbesman, S.H. Strogatz, M.S. Vitevitch, Comparative analysis of networks of phonologically similar words in English and Spanish, Entropy 12 (3) (2010) 327–337.
[18] K.Y. Chan, M.S. Vitevitch, The influence of the phonological neighborhood clustering-coefficient on spoken word recognition, J. Exp. Psychol. Human Perception Perform. 35 (2009) 1934–1949.
[19] R. Ferrer i Cancho, R.V. Solé, R. Köhler, Patterns in syntactic dependency networks, Phys. Rev. E 69 (2004) 051915.
[20] H. Liu, The complexity of Chinese dependency syntactic networks, Physica A 387 (2008) 3048–3058.
[21] H. Liu, F. Hu, What role does syntax play in a language network? EPL 83 (2008) 18002.
[22] H. Liu, W. Li, Language clusters based on linguistic complex networks, Chinese Sci. Bull. 55 (30) (2010) 3458–3465.
[23] B. Corominas-Murtra, S. Valverde, R.V. Solé, The ontogeny of scale-free syntax networks: phase transitions in early language acquisition, Adv. Complex Syst. 12 (2009) 371–392.
[24] B. Corominas-Murtra, S. Valverde, R.V. Solé, Emergence of scale-free syntax networks, in: E. Nolfi, M. Mirolli (Eds.), Evolution of Communication and Language in Embodied Agents, Springer, 2010, pp. 83–99.
[25] R. Čech, J. Macutek, Word form and lemma syntactic dependency networks in Czech: a comparative study, Glottometrics 19 (2009) 85–98.

[26] M. Steyvers, J.B. Tenenbaum, The large-scale structure of semantic networks: statistical analyses and a model of semantic growth, Cogn. Sci. 29 (2005) 41–78.
[27] H. Liu, Statistical properties of chinese semantic networks, Chinese Sci. Bull. 54 (16) (2009) 2781–2785.
[28] J. Borge-Holthoefer, A. Arenas, Semantic networks: structure and dynamics, Entropy 12 (5) (2010) 1264–1302.
[29] T.T. Hills, M. Maouene, J. Maouene, A. Sheya, L Smith, Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? Psychol. Sci. 20, 25 (2009) 729–739.
[30] P. Erdös, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960) 17–61.
[31] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[32] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2003) 167–256.
[33] R. Albert, H. Jeong, A.-L. Barabasi, Error and attack tolerance of complex networks, Nature 406 (2000) 378–382.
[34] P. Menzerath, Die Architektonik des deutschen Wortschatzes, Dümmler, Bonn, 1954, p100.
[35] G. Altmann, Prolegomena to Menzerath's law, Glottometrika 2 (1980) 1–10.
[36] A. Fenk, G. Fenk-Oczlon, L. Fenk, Syllable complexity as a function of word complexity, in: V. Solovyev, V. Goldberg, V. Polyakov (Eds.), The VIIIth International Conference "Cognitive Modeling in Linguistics" (Text Processing and Cognitive Technologies, 1), Kazan State University Press, Kazan, 2006, pp. 324–333.
[37] G.K. Zipf, Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology, Addison-Wesley, Cambridge, MA, 1949.
[38] R. Ferrer i Cancho, R.V. Solé, Least effort and the origins of scaling in human language, Proc. Natl. Acad. Sci. USA 100 (2003) 788–791.