# The Study of Phonological Neighborhoods in Chinese L1 and L2 Speech Production

Tongtong Xie
School of Foreign Languages, Shenzhen University, 3688 Nanhai Ave, Nanshan, Shenzhen, Guangdong 518060, P.R. China

Haiying Ye
School of Foreign Languages, Shenzhen University, 3688 Nanhai Ave, Nanshan, Shenzhen, Guangdong 518060, P.R. China

Hongyan Wang
School of Foreign Languages, Shenzhen University, 3688 Nanhai Ave, Nanshan, Shenzhen, Guangdong 518060, P.R. China
wanghongyan0069@hotmail.com

Jeroen van de Weijer
School of Foreign Languages, Shenzhen University, 3688 Nanhai Ave, Nanshan, Shenzhen, Guangdong 518060, P.R. China

## ABSTRACT

The tongue twister paradigm was used to compare numbers and types of errors of native and non-native speakers of Chinese when producing tongue twisters. The stimuli consisted of 106 quadruples, 32 of which were transliterated from English tongue twisters, 26 of which were vocalic twisters, and 48 of which were consonant twisters. Both consonant and vowel errors were investigated (but not tone) and errors were classified as caused by either preceding or following linguistic forms (or as caused by both or neither). To enhance errors, we requested participants to use a speech rate that was 20% faster than the normal rate. Four native Mandarin Chinese speakers and six foreign learners of Chinese read the tongue twisters aloud, repeating each one four times in a slide. The four native Mandarin Chinese speakers made a total of 606 errors, and the non-native speakers produced 3970 errors. The results show a clear difference between L1 and L2 speakers and a relation between years of learning Chinese and total number of errors.

## CCS CONCEPTS

• **Applied computing** → Education; Computer-assisted instruction.

## KEYWORDS

Tongue twisters, Chinese, native vs. non-native speech, error analysis

**ACM Reference Format:**
Tongtong Xie, Haiying Ye, Hongyan Wang, and Jeroen van de Weijer. 2020. The Study of Phonological Neighborhoods in Chinese L1 and L2 Speech Production. In *2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence (ACAI 2020), December 24–26, 2020, Sanya, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3446132.3446135


*ACAI 2020, December 24–26, 2020, Sanya, China*

ACM ISBN 978-1-4503-8811-5/20/12…$15.00
https://doi.org/10.1145/3446132.3446135

## 1 INTRODUCTION

Speech errors form an important source of evidence on how language is acquired, stored, produced and perceived. For instance, children's errors in morphology (*I breaked it*) show that they generalize regular rules to irregular forms. Spoonerisms (*our dizzy bean* for *our busy dean*) show that individual segments play a role in speech planning and inform theories of syllable structure (see e.g. [1], [2], [3], [4]). Most of these "error corpora" contain incidental collections of speech errors and can therefore not be regarded as balanced speech data. The recently developed tongue twister paradigm ([5, 6]) has added to the error-based approach in a constructive way, because it allows researchers to collect speech errors systematically and in a controlled experimental setting. In this methodology, subjects are asked to repeat "twistable" phrases (e.g. sequences of words containing the same or similar vowels, or consonants, or tones – all can be exactly controlled), often under time pressure, to answer specific questions about the organization of the lexicon (e.g. [7, 8]), and the speech planning process, especially with respect to the production stage (e.g. [9]). Questions can be raised such as: do speakers plan ahead when reading words, and does this differ between languages (e.g. with different orthographic systems such as English and Chinese), if they make mistakes do they tend to replace segments (vowels and consonants) with segments that came before (showing that they keep already pronounced segments in a kind of "memory buffer") or do they tend to replace them with segments that are still to come (showing that they have already "preselected" such segments for quick access); or both (if so, which strategy is more prevalent?). Other questions are whether segments tend to be replaced by segments that are similar (e.g. stops by stops, fricatives by fricatives, or having a similar place of articulation, and so on, showing the role of phonological "natural classes" in speech planning), whether native speakers make the same kinds of errors as second language learners (i.e. if there is an interference effect even if these errors) and if bilinguals "twist" in the same way as monolinguals.

In recent years this research paradigm has produced interesting results, also for Chinese (e.g. [10], [11]), but much research remains to be done. In this project we specifically aim to answer questions about effects of native language, the direction of twisting and the role of natural classes. To achieve this aim, we recorded both Chinese native speakers, using the tongue twister paradigm, and foreign learners of Chinese, and analysed their data statistically. In this way we increased understanding of how speech errors arise, which may help future treatment and training applications.

We were particularly interested in the question if native speakers and non-native speakers make similar numbers of errors and whether they display similar types of errors, e.g. with respect to consonants and vowels, and in the direction of errors (either anticipatory or perseverative (or both)). We excluded tonal errors from the analysis, partly because of the fact that tonal L2 acquisition is notoriously hard, so many errors were expected anyway. As to the number of errors, on the one hand it might be expected that non-native speakers make more errors than native ones on general grounds, but on the other hand they might pay more conscious attention to (especially written) text than native speakers. As to whether they would make more errors in consonants than with vowels, or vice versa, we had no *a priori* expectation.

## 2 EXPERIMENT

To test these hypotheses, we carried out an explorative production experiment in which native and non-native Chinese speakers produced numerous tongue twisters in Chinese, adopting the tongue twister paradigm. We analyzed the results statistically.

### 2.1 *Materials*

We used the general research set-up of the tongue twister paradigm [5]. The tongue twisters all consisted of four Chinese syllables. Half of these had consonants that were potential targets for twisting (example: 架票跳恰, pinyin equivalent: *jia piao tiao qia*) and half had vowels that were targets for twisting (example: 谬念妙泞, *miu nian miao ning*). Both types were construed with similar and with dissimilar vowels and consonants, respectively, and in two patterns (ABBA and ABAB, see below). Besides the target stimuli, there was also a control set (example: 地嗽套破, *di sou tao po*). The stimuli consisted altogether of 106 quadruples, 32 of which were transliterated from English tongue twisters (ibid.) (e.g. English "cub time date sin" was rendered as Chinese "快太对四" (pinyin equivalent: *kuai tai dui si*). Like in the English examples, these were not possible sentences, but just sequences of words. Vowel and consonant similarity and dissimilarity were based on Chinese phoneme dendrograms [12]. In order to disregard any tonal effects, tone was constant on all Chinese characters (always tone 4). All characters were among the most frequent in the *Modern Chinese Word Frequency Dictionary*[13]. For the vocalic twisters, we selected two consonants and four similar and four dissimilar in vowels to make sure every consonant could be matched at least once (e.g. 谬念妙泞, *miu nian miao ning*). We selected /m/ and /n/ as consonants and /iu/, /ia/, /iao/, and /i/ as similar vowels. Conversely, in the consonant group, we selected two vowels with four similar (or dissimilar, in the corresponding dissimilar group) consonants (e.g. 烂让汉放, *lan rang han fang*). In cases where particular consonants could not be combined with particular vowels (due to phonotactic constraints in Chinese), consonants in the tongue twister were replaced with /j/ (e.g.月窃叶略, *yue qie ye lüe*).

After selecting the characters, all stimuli were divided into either an ABBA or an ABAB format, following [5], [11], to test the direction of tongue twisting. For example, the tongue twister谬念妙泞(*miu nian miao ning*) presents the target consonants /m/ and /n/ in ABAB order, while in 派跳退破(*pai tiao tui po*) the target consonants /p/ and /t/ are presented in ABBA order. Similarly, for vowels, in the tongue twister 六跳吊秀(*liu tiao diao xiu*), the target vowels /iu/ and /iao/ appear in ABBA order, and in 个末设破(*ge mo she po*), the vowels /e/ and /o/ appear in ABAB order.

**Table 1:** Participants (N=native Chinese speaker; W=L2 Chinese speaker

| | L1 | gender | age | years of Chinese (for L2 speakers) |
|---|---|---|---|---|
| N1 | Chinese | F | 19 | |
| N2 | Chinese | F | 19 | |
| N3 | Chinese | M | 19 | |
| N4 | Chinese | M | 19 | |
| W1 | Urdu | M | 33 | 3 |
| W2 | Russian | F | 25 | 5 |
| W3 | English | F | 62 | 10 |
| W4 | Dutch | F | 52 | 20 |
| W5 | English | M | 48 | 14 |
| W6 | English | M | 33 | 13 |

### 2.2 *Participants*

We recruited four native speakers of Chinese who were all students at Shenzhen University (two male, two female, average age 19). We also recruited six non-native speakers of Chinese (three males, three females, average age 42) who worked as teachers at the same university and had different lengths of experience with Chinese (average: 11 years). An overview of all speakers is provided in Table 1

### 2.3 *Procedure*

Subjects attended the experiment individually in a quiet room of the Phonetics Laboratory of the School of Foreign Languages, Shenzhen University. Items were presented on PowerPoint slides on a Tsinghua Tongfang desktop computer monitor. For the Chinese L2 speakers, word quadruples were displayed in pinyin in bold, black, Arial 60-point lowercase font on a white background. For the L1 speakers, character quadruples were presented in bold, black Heiti 60-point font on a white background. Subjects were asked to recite each slide four times, after a training session to become accustomed to the task. In order to increase time pressure on the subjects, to induce them to make errors, slides were presented for a limited period of time, which we calculated at 20% faster than spontaneous speech rate [5, 14, 15]. For the slides in pinyin, time was set for 5.33 seconds per slide and for the slides with Chinese characters, the time for each slide was set for 5 seconds. For all subjects, the interslide interval was 3-5 seconds (the interslide interval decreased from 5 seconds in the first part of the experiment to 3 seconds in the last part, to keep up the time pressure).

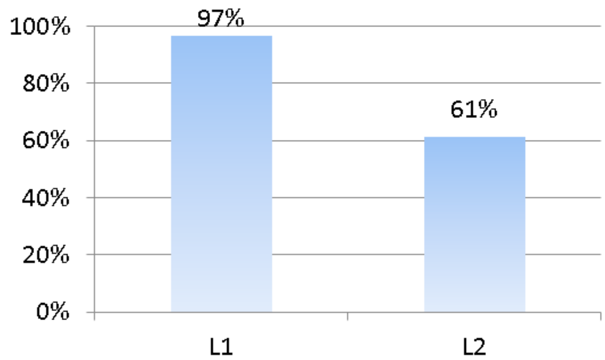


**Figure 1: Percentage of input stimuli realised, for L1 speakers and L2 speakers**

Participants were asked to read the tongue twisters aloud four times on each slide. After the first click, the target phrase would appear on the screen automatically. After finishing each item, the subjects clicked again and went to the next test item. Most subjects could pronounce the tongue twister successfully for four times, while some claimed the speed was too fast for them. The slides were arranged in order of Control, Alliterating Similar, and Alliterating Dissimilar, in order to familiarize the subjects with the task. If a subject pronounced a tongue twister three times instead of four, one missing value was scored. At least two syllables must be present to constitute a data point. All in all, the four native speakers pronounced 106 tongue twisters four times ($4 \times 106 \times 4 = 1{,}696$ potential data points) and the six L2 speakers pronounced 106 tongue twisters four times ($6 \times 106 \times 4 = 2{,}544$ potential data points), where a 'data point' is one fully or partially pronounced tongue twister.

All responses were recorded by a Sennheiser headset USB microphone connected to a professional TASCAM DR-05 recorder. After each session, the recordings were stored on computer disk for further phonetic and statistical analysis.

## 2.4 Results

Of 1,696 potential data, the L1 group produced 1,640 data points (either a complete tongue twister or at least two syllables out of four), or 97%. The L2 group produced 1,555 data points out of 2,544 potential data points (61%). Thus, the L2 groups had more difficulty with the tongue twisters than the L1 group, as Figure 1 shows:

Second, we turn to the number of total errors. Table 2 and Table 3 show the number of consonantal and vocalic errors (and tonal) for both groups.

Consonantal errors are not greatly different between the two groups. Vocalic errors about three times as frequent in the L2 speakers than in the L1 speakers. The biggest difference, however, lies in the tones.

The following figure shows the total number of errors compared to the years of active Chinese for the L2 group (see Table 1).

So the person with least experience with Chinese showed the largest number of errors, and the person with most experience

**Table 2:** Number of Consonant, Vowel and Tone errors, for L1speakers

| L1 | Total C | Total V | Total T | Total all |
|---|---|---|---|---|
| Speaker1 | 105 | 55 | 0 | 160 |
| Speaker2 | 107 | 32 | 9 | 148 |
| Speaker3 | 83 | 39 | 8 | 130 |
| Speaker4 | 148 | 16 | 4 | 168 |
| Sum | 443 | 142 | 21 | 606 |
| Percentage | 27% | 9% | 1% | 37% |
| Average | 110.75 | 35.5 | 5.25 | 151.5 |

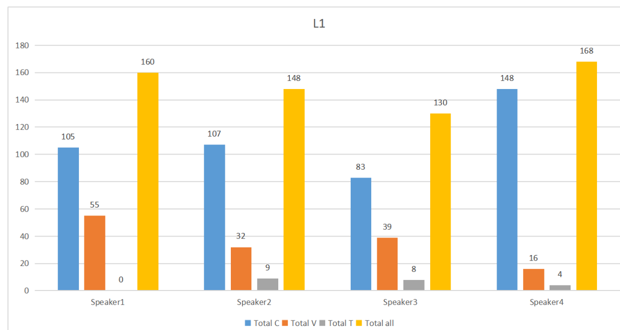


**Figure 2: Graph of Consonant, Vowel and Tone errors, for L1speakers**

**Table 3:** Number of Consonant, Vowel and Tone errors for L2 Speakers

| L2 | Total C | Total V | Total T | Total all |
|---|---|---|---|---|
| Speaker5 | 119 | 128 | 1043 | 1290 |
| Speaker6 | 116 | 47 | 432 | 595 |
| Speaker7 | 89 | 99 | 216 | 404 |
| Speaker8 | 7 | 10 | 6 | 23 |
| Speaker9 | 192 | 136 | 651 | 979 |
| Speaker10 | 70 | 41 | 568 | 679 |
| Sum | 593 | 461 | 2916 | 3970 |
| Percentage | 38% | 30% | 188% | 255% |
| Average | 98.8 | 76.8 | 486 | 661.7 |

showed the smallest number of errors. W4 and W5 are not regular, however.

Recall that errors could be caused by segmental material to the left, to the right, both sides, or neither (or that could be a nasal error). Table 4 presents the overall results, first in terms of direction (left, right, both, neither, or a nasal error; top panel) and then in terms of pattern (ABBA vs. ABAB, bottom panel), each divided for L1 and L2 speakers:

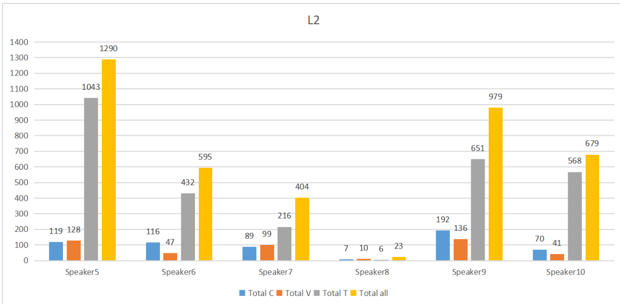


**Figure 3: Graph of Consonant, Vowel and Tone errors, for L2 speakers**

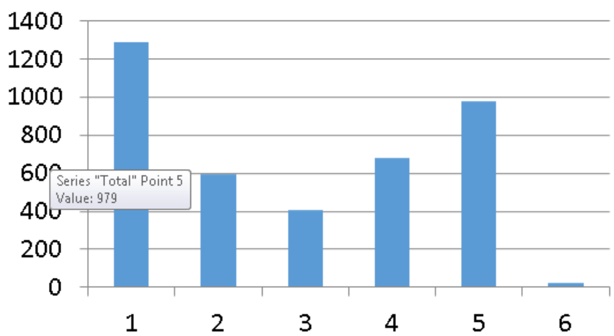


**Figure 4: Bar diagram for years of active Chinese by total number of errors**

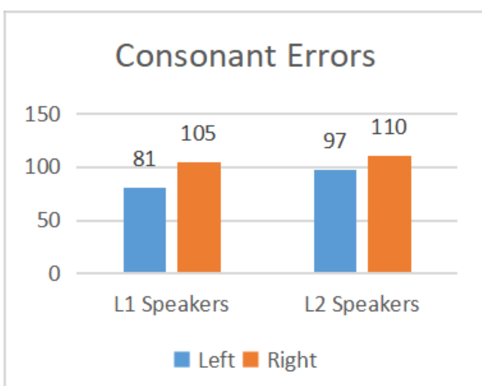


**Figure 5: Graph of left and right errors in Consonant errors, for L1 and L2 speakers**

## 3 DISCUSSION AND CONCLUSION

The results presented above clearly show that tongue twisters are more difficult for L2 speakers than for native speakers. L2 speakers have greater difficulty completing the tongue twisters and make more errors (Figure 1, Table 2 and Table 3). The number of errors is roughly equally divided across consonants and vowels, but tonal errors are even much more common in the L2 speakers. There was

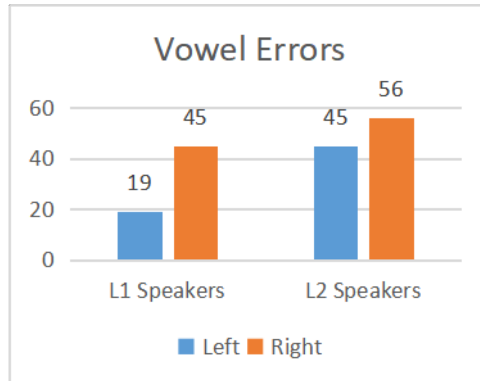


**Figure 6: Graph of left and right errors in Vowel errors, for L1 and L2 speakers**

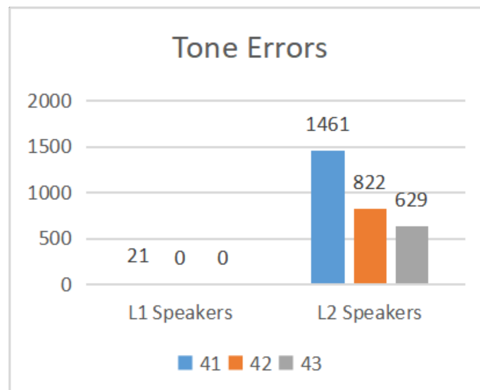


**Figure 7: Graph of 41, 42 and 43 errors in Tone errors, for L1 and L2 speakers**

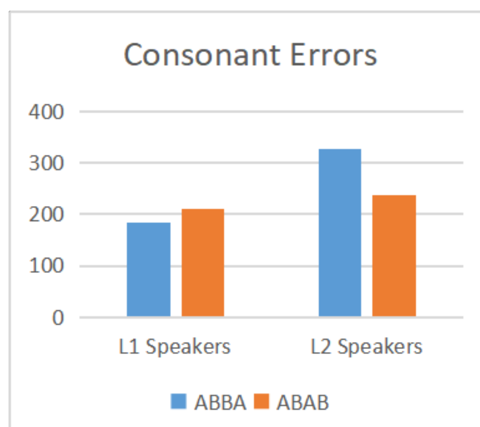


**Figure 8: Graph of ABBA and ABAB errors in Consonant errors, for L1 and L2 speakers**

**Table 4:** Number and Graphs of segment and tone errors for L1 and L2

| | | |
|---|---|---|
| Consonant Errors | | |
| L1 Speakers | Left | 81 |
| | Right | 105 |
| | Both | 102 |
| | Neither | 129 |
| | Nasal | 26 |
| L2 Speakers | Left | 97 |
| | Right | 110 |
| | Both | 108 |
| | Neither | 196 |
| | Nasal | 82 |
| Vowel Errors | | |
| L1 Speakers | Left | 19 |
| | Right | 45 |
| | Both | 18 |
| | Neither | 60 |
| L2 Speakers | Left | 45 |
| | Right | 56 |
| | Both | 27 |
| | Neither | 333 |
| Tone Errors | | |
| L1 Speakers | 41 | 21 |
| | 42 | 0 |
| | 43 | 0 |
| L2 Speakers | 41 | 1461 |
| | 42 | 822 |
| | 43 | 629 |
| Consonant Errors | | |
| L1 Speakers | ABBA | 185 |
| | ABAB | 212 |
| L2 Speakers | ABBA | 327 |
| | ABAB | 238 |
| Vowel Errors | | |
| L1 Speakers | ABBA | 57 |
| | ABAB | 42 |
| L2 Speakers | ABBA | 170 |
| | ABAB | 158 |
| Tone Errors | | |
| L1 Speakers | ABBA | 9 |
| | ABAB | 19 |
| L2 Speakers | ABBA | 917 |
| | ABAB | 1054 |

a clear relation between years of active Chinese and total number of errors (Figure 4).

L2 speakers of Chinese had a diverse number of language backgrounds (Table 1). With larger numbers of L2 speakers it might be possible to relate native language to different kinds of errors. For instance, languages differ in consonant and vowel inventory, which may affect their performance on L2 tongue twisters.

In future work, we will present further results as to kinds of errors, e.g. in relation to ABBA or ABAB format, and extend this research paradigm in order to shed further light on how speech errors can be related to linguistic processing.

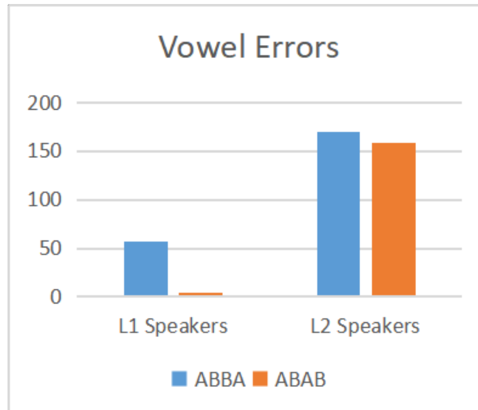


**Figure 9: Graph of ABBA and ABAB errors in Vowel errors, for L1 and L2 speakers**

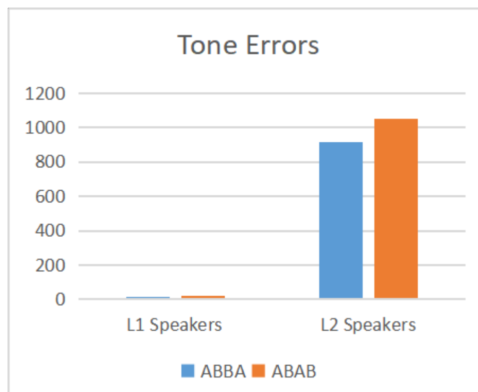


**Figure 10: Graph of ABBA and ABAB errors in Tone errors, for L1 and L2 speakers**

## REFERENCES


[1] MacKay, D.G.: 'Spoonerisms: The structure of errors in the serial order of speech', Neuropsychologia, 1970, 8, (3), pp. 323-350

[2] Shattuck-Hufnagel, S.: 'The representation of phonological information during speech production planning: Evidence from vowel errors in spontaneous speech', PY, 1986, 3, pp. 117-149

[3] Slis, A.W.: 'Articulatory variability and speech errors: An overview', Toronto Working Papers in Linguistics, 2018, 40, (1), pp. 1-16

[4] Ran, Q., and Yu, S.: 'Characteristics and patterns of Chinese speech errors: Analysis and exploration into the error-item dataset from 25 L1 backgrounds', Chinese Teaching in the World, 2019, 33, (3), pp. 417-432

[5] Wilshire, C.E.: 'The "tongue twister" paradigm as a technique for studying phonological encoding', LSp, 1999, 42, (1), pp. 57-82

[6] Wilshire, C.E.: 'Serial order in phonological encoding: an exploration of the 'word onset effect' using laboratory-induced errors', Cog, 1998, 68, (2), pp. 143-166

[7] Vitevitch, M.S.: 'Tongue twisters reveal neighborhood density effects in speech production', Research on spoken language processing, progress report, 1996, 21, pp. 545-556

[8] Vitevitch, M.S.: 'The influence of phonological similarity neighborhoods on speech production', Journal of Experimental Psychology: Learning, Memory, and Cognition, 2002, 28, (4), pp. 735

[9] Corley, M., Brocklehurst, P.H., and Moat, H.S.: 'Error biases in inner and overt speech: Evidence from tongue twisters', Journal of Experimental Psychology: Learning, Memory and Cognition, 2011, 37, (1), pp. 162-175

[10] Zhang, S., and Perfetti, C.A.: 'The tongue-twister effect in reading Chinese', Journal of Experimental Psychology: Learning, Memory, and Cognition, 1993, 19, (5), pp. 1082-1093

[11] Kember, H., Croot, K., and Patrick, E.: 'Phonological encoding in Mandarin Chinese: Evidence from tongue twisters', LSp, 2014, 58, (4), pp. 417-440

[12] Zhang, J.: 'The distinctive features for Standard Chinese ( *putonghua*)', Acta Acustica, 2005, 30, (6), pp. 506-514

[13] Chinese Academy of Social Sciences: 'Modern Chinese word frequency dictionary' (Chinese Press, 1988)

[14] Croot, K., Au, C., and Harper, A.: 'Prosodic structure and tongue twister errors', in Fougeron, C., Kühnert, B., D'Imperio, M., and Vallée, N. (Eds.): 'Laboratory Phonology' (Walter de Gruyter, 2010), pp. 433-461

[15] Fónagy, I., and Magdics, K.: 'Speed of utterance in phrases of different lengths', LSp, 1960, 3, (4), pp. 179-192