*Concentric: Studies in Linguistics*
39.1 (May 2013): 1-21

# The Effects of Contextual Availability and Phonetic Similarity on Speech Errors*

Joyce H.-C. Liu
*Kai Nan University*

This study explores the effects of contextual availability and phonetic similarity on speech errors of Taiwan Southern Min. The results show that most consonant errors are attributed to the influence of surrounding segments. The number of error tokens and the distance of contextual availability reveal an inverse relationship. The limited distance of 7 syllables as a domain of the contextual-availability effect is of significance concerning the linguistic unit in language processing. On the other hand, the results show that the number of error tokens is directly proportional to the degree of phonetic similarity shared by the target and the intruder. In summary, the results support the findings shown in previous studies of other languages, implying some universal properties of speech errors. Furthermore, this study shows that the effect of phonetic similarity may vary with the featural system used as an index of phonetic similarity.

Key words: speech errors, consonant, contextual availability, phonetic similarity, Taiwan Southern Min

## 1. Introduction

Speech errors provide a different source of data in psychological research to explore the process of language and the model of language production (Dell & Reich 1977, Bock 1982, Levelt 1989), and in linguistic research to assess the psychological reality of phonological components as well as the syllable structure (Fromkin 1973, 1989, Cutler 1982, Shattuck-Hufnagel 1986, Berg 1987). Studies show that speech errors do not occur at random but are constrained or influenced by some factors (Cohen 1973, Nooteboom 1973, Stemberger 1989, Jaeger 2005, Wan 2007). Contextual availability and similarity effects are two of the factors contributing to the occurrence of phonological errors that have been studied in different languages with consistent results. The effect of contextual availability shows that most phonological errors are attributed to the influence of a potential intruder available in the context within a distance (Cohen 1973, Nooteboom 1973, Wan 2007, Liu 2009). The distance domain between the target, i.e. the intended unit, and the intruder is within 7 syllables most of the time and usually within the same clause. The similarity effect is a cover term for several structural aspects, including phonological similarity, stress similarity, phonemic similarity, and phonetic similarity. These structural similarities between the target and the intruder in speech errors are found to facilitate the occurrence of speech

* The content of this paper is revised from part of my doctoral dissertation (Liu 2009) with different analyses on the data. In addition, part of this paper was presented in the 12th conference on Min languages held in Academia Sinica, Taipei.

errors (Boomer & Laver 1968, MacKay 1970, Fay & Cutler 1977, Shattuck-Hufnagel 1985, Jaeger 2005). Most of the target and the intruding segments share phonological similarity in terms of the syllable position. Onset consonants interact with onset consonants, and coda consonants with coda consonants. The two segments involved in errors also share phonetic similarity in terms of segmental features. They usually differ in only one or two features. Moreover, the target and intruding syllables usually belong to the same stress position and bear phoneme similarity by sharing common components such as onset consonant, vowel or rime.

In this study, we explore the effects of contextual availability and phonetic similarity on consonant errors of Taiwan Southern Min (hereafter TSM). The effect of contextual availability is assessed in three ways:

(1) Do most of the speech errors occur with a potential intruder in the context?
(2) Given a contextual segmental error, how far could the intruder be away from the target syllable? Is there a domain of contextual influence in terms of syllables?
(3) Is there any relation between the distance and error tokens? Does the distance between the target and the intruder affect the error rate?

Concerning the similarity effect, though several similarity effects have been discussed in previous studies, TSM is not a good candidate to investigate the effect of phonological similarity, stress similarity, nor phoneme similarity on speech errors due to its phonological structure. On the one hand, unlike intonation languages such as English, there are no obvious lexical stress patterns in TSM as it is a tone language. On the other hand, unlike many languages allowing consonant clusters, the syllable structure of TSM is relatively simple. It can be represented by a formula (C)(G)V(E), in which C stands for a consonant, G for a glide, V for a vowel, and the final component ‘E’ for either a coda consonant or a glide. All components except the vowel are optional. The simplicity of the TSM syllable structure makes it easy for the target and the intruding syllables in speech errors to have the same syllable structure as well as common phonemes by chance. Therefore, we only examine the effect of phonetic similarity in this study.

The phonetic similarity between two segments is usually measured by segmental features proposed in linguistic theories. Two segments sharing more features in common indicate a higher phonetic similarity between them. However, there are several featural systems proposed in linguistic study. These systems differ slightly on the number of composed features, and the numerical discrepancy might lead to confusion on interpretation and on cross-study comparison. For example, one study adopts a system composed of five features as an index of phonetic similarity, while another uses a system composed of four features. Two segments A and B share four common phonetic features based on the five featural system. The other two segments

C and D share three common features based on the four featural system. The two segments A and B are no more similar to each other than the pair C and D, though the former shares one more common feature. In terms of feature differences, both pairs differ by one feature. In other words, the two pairs of segments actually have the same degree of phonetic similarity. Hence, for the clarity of interpretation and ease of cross-study or cross-system comparison, the phonetic similarity is conventionally evaluated by counting the number of features two segments differ by. The lower the number of feature differences, the higher the phonetic similarity. In this study, we also examine the effect of phonetic similarity by measuring feature differences shared by two segments involved in errors.

An issue which is fundamentally related to the phonetic similarity but seldom discussed in literature is the appropriateness of a featural system adopted as an index of phonetic similarity. As mentioned above, different featural systems vary on the number of composed features. The place of articulation of a consonant can be represented by two features 'anterior' and 'coronal' (Chomsky & Halle 1968) or a cover term 'place of articulation' with multi-values (Van den Broecke & Goldstein 1980, Jaeger 1992, Wan 1999). It is also of our interest to investigate if the adoption of different featural systems biases the results of phonetic similarity.

In summary, three different featural systems are adopted to assess the effect of phonetic similarity by answering the following questions:

(1) Given two segments involved in errors, do these two segments share phonetic similarity? If yes, how similar are they? In other words, how many features do they usually differ by?
(2) Is there any correlation between phonetic similarity and error tokens?
(3) Are there any differences among the featural systems used to index similarity? Is there any system better than the others as an index of phonetic similarity?

## 2. The consonantal inventory of TSM

There are fifteen consonantal phonemes in TSM, as shown in Table 1. They are called "the fifteen sounds" in traditional phonology books (Yang 1995). There are voicing and aspiration contrasts among stops and affricates. Voiced stops [b, l, g] and the corresponding nasal stops [m, n, ŋ] are allophones in complementary distribution. Oral stops only occur with oral vowels and nasal stops appear with nasal vowels. Dental fricatives [s] and affricates [ts] and [$ts^h$] will undergo palatalization before the front high vowel [i].

All consonants can occur in the syllable-initial position with a distinctive function except for the glottal stop [ʔ]. The glottal stop is regarded as "zero onset", and occurs in the syllable-initial position optionally, varying among speakers (Yang 1995). Compared with initial consonants, there is a much stricter restriction on coda consonants. Only unaspirated stops, nasals, and the glottal stop can occur in the syllable-final position.[1]

**Table 1. The consonantal inventory of TSM**

| | | | Labial | Dental | Velar | Glottal |
|---|---|---|---|---|---|---|
| Plosive | Voiced | | b (m) | l* (n) | g (ŋ) | |
| | Voiceless | Unaspirated | p | t | k | ʔ |
| | | Aspirated | $p^h$ | $t^h$ | $k^h$ | |
| Affricate | Voiced | | | dz** | | |
| | Voiceless | Unaspirated | | ts (tɕ) | | |
| | | Aspirated | | $ts^h$ ($tɕ^h$) | | |
| Fricative | Voiceless | | | s (ɕ) | | h |

* The phonetic property of the voiced dental stop is somewhere between [l] and [d]. It is used to be labelled as /l/ in traditional Chinese linguistics (Yang 1995).
** The voiced affricate [dz] is pronounced as a fricative [z] in some dialects, and is changed to [l] or [g] in some other dialects (Yang 1995).

## 3. Data

In this study, speech errors refer to one-time errors occurring in speech production which are involuntary deviations from the speaker's intention in regards to phonological, lexical, or grammatical aspects (Sturtevant 1947, Boomer & Laver 1968, Jaeger 2005). The data source is an on-going TSM corpus containing recordings of spontaneous speech collected from fifteen radio programs.[2] The subjects comprise more than one hundred people, including hosts and hostesses of the programs, invited guests, and call-in audience. A total of 2600 speech errors were collected from approximately 1200 recording files, with a total recording time of almost 500 hours. About ten percent of the data was randomly selected to be double-checked by three trained linguists.[3] The accuracy rate of the transcription is over 85%, which ensures the high reliability of the transcription.

Data used for analysis in this study only include phonological errors involving substitution of a single onset consonant or exchange of two onset consonants. Errors involving coda consonants are excluded in this study due to the limited number and the controversial phonemic status of coda consonants. Initial-consonant errors with

[1] The phoneme status and the phonetic properties of coda consonants in TSM are still controversial.
[2] Taiwan Southern Min spoken corpus is constructed and maintained by the Institute of Linguistics at National Chung Cheng University.
[3] The present author would like to express her gratitude to Professor Jane Tsay and three graduate students in the Linguistics program, Houyou Pan, Yunwei Li, and Zhequan Zhang, for their assistance in double-checking the data.

other possible alternative analysis are also excluded from analysis. There are 273 consonant errors in total included in this study.

## 4. Analysis 1: Contextual availability and the domain of the target-intruder distance

### 4.1 Contextual availability

It is reported that most phonological errors are contextual due to the influence of the surrounding linguistic unit and that the potential intruder is usually located within a certain domain, i.e. 7 syllables, from the target syllable (Cohen 1973, Nooteboom 1973, Levelt 1989). As the domain of the target-intruder distance (henceforth the T-I distance) is also an issue in question, in order not to be biased by the previous studies, we set a broader criterion for a potential intruder of contextual errors. Errors with a potential intruder located within 12 syllables from the target syllable and within one utterance[4] in principle are counted as contextual errors. Errors without a potential intruder in the utterance are treated as non-contextual errors.

The T-I distance is measured by syllables. If the intruder is located in a syllable right adjacent to the target syllable, the T-I distance is counted as one syllable. If there is more than one potential intruder in an utterance, a minimal-distance principle is adopted. Hence, the shortest distance is counted. Take (1) for illustration.

(1) Intended: **t**am33<u>l</u>uan33ai21 e33 <u>l</u>am33 <u>l</u>u55 piŋ33iu53
Error: **<u>l</u>**am33
'The couples who are falling in love'

In (1), the target segment [t], marked with a boldfaced font in the intended utterance, is mispronounced as [l], which is boldfaced and underlined in the erroneous utterance. There are three potential intruders in the intended utterance, all marked with an underline: onset consonant [l] of the syllable [luan33], onset of the syllable [lam33] and onset of the syllable [lu55]. Based on the minimal-distance principle, the syllable [luan33] is selected for the measurement of the T-I distance as it is the nearest syllable from the target. Therefore the T-I distance of this error token is 1 syllable.

Among the 273 tokens, 259 of them (94.87%) are labelled as contextual errors for at least one potential intruder can be found in the context. Only 14 tokens (5.13%) with no potential intruding segment found in the utterance are treated as non-contextual errors. Although we set up a criterion of 12 syllables from the target

[4] There are some errors with the target syllable in the beginning or final position of an utterance. In these situations, a potential intruder is allowed to cross the clause boundary.

syllable as a domain of the T-I distance, there is no error with a T-I distance of more than 7 syllables. All contextual errors have a potential intruder within 7 syllables from the target syllable. Hence, the result of contextual availability is consistent with previous studies in that most of the phonological errors are contextual errors with a potential intruder available in the utterance.

Nevertheless, there is also a discrepancy on the percentage of contextual errors between our study and previous studies. As reported in the study of English speech errors (Shattuck-Hufnagel & Klatt 1979), contextual errors account for about 70 percent of phonological errors. The percentage of contextual errors in TSM is much higher than that in English. We noticed that it seems frequent for a segmental error to contain more than one potential intruder in the utterance when measuring the T-I distance. This is confirmed with an examination of error tokens based on the number of potential intruder within the domain of 7 syllables from the target, as shown in Table 2.

**Table 2. The tokens of errors based on the number of potential intruder**

| N. of potential intruder | 1 | 2 | 3 | 4 | Total |
|---|---|---|---|---|---|
| Tokens | 166 | 76 | 14 | 3 | 259 |
| % | 64.09 | 29.34 | 5.41 | 1.16 | 100 |

It is clear in Table 2 that in more than one-third of the contextual errors (35.91%), at least two potential intruders can be found in the utterance. In some tokens, there are four potential intruders found in the utterance, as an example shown in (2)

(2) Intended: li55 **ts**ui53 i55kiŋ33 <u>ts</u>$^h$ioŋ33 <u>ts</u>$^h$iu53 <u>ts</u>$^h$ioŋ33 <u>ts</u>$^h$iŋ33k$^h$i21 a0
Error: **<u>ts$^h$</u>**ui53
'You have cleaned your hands with the water.'

In (2), the target segment [ts] is mispronounced as [ts$^h$]. There are 4 syllables with an onset consonant [ts$^h$] in the utterance, and all of them are no farther than 12 syllables from the target syllable [tsui53] 'water'. Hence, the consonant [ts$^h$] in all 4 syllables are treated as potential intruders.

As there is no information provided in Shattuck-Hufnagel & Klatt's study concerning the error tokens with multiple potential intruders, there is no way to make a comparison. Nevertheless, the high percentage of errors with more than one potential intruder in TSM probably provides an explanation for the higher percentage of contextual errors in TSM than in English. A further discussion on this issue will be made in Section 6.

### 4.2 The domain of the T-I distance

The distance between the target and the intruder involved in contextual errors are measured to identify the possible domain of availability. The results are shown in Table 3 and Figure 1.

**Table 3. Distribution of contextual error tokens based on the T-I distance**

| Distances (syl.) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | >7 | Total |
|---|---|---|---|---|---|---|---|---|---|
| Tokens | 152 | 50 | 27 | 14 | 6 | 5 | 5 | 0 | 259 |
| % | 58.69 | 19.31 | 10.42 | 5.41 | 2.32 | 1.93 | 1.93 | 0 | 100 |
| Cumulative % | 58.69 | 78 | 88.42 | 93.83 | 96.15 | 98.08 | 100 | 100 | 100 |

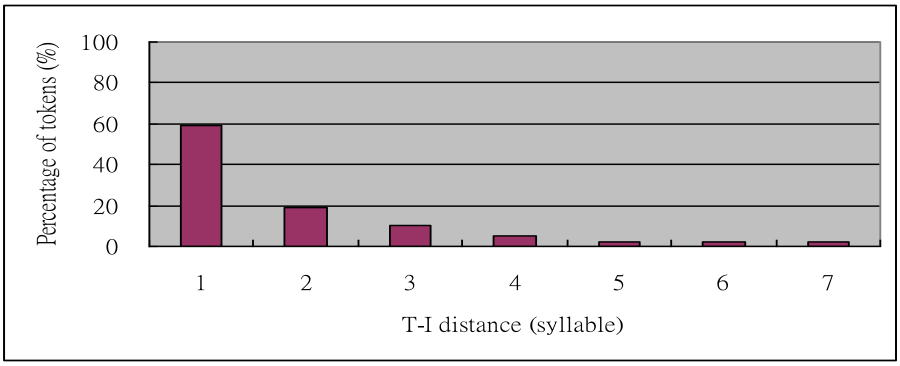


**Figure 1. Distribution of contextual error tokens based on the T-I distance**

The distribution pattern shown above reveals a significant characteristic. The number of error tokens is inversely related to the T-I distance. A linear correlation analysis shows a significantly negative correlation between the T-I distance and the error tokens ($p < .001$). This characteristic can be interpreted in two ways. One is that the largest proportion of tokens (N=152, 58.7%) involve errors with a T-I distance of 1 syllable. That is, more than half of the errors are attributed to the influence of adjacent syllables. The average T-I distance is 1.87 syllables. The other is that although the possible T-I domain contains a distance of 7 syllables from the target, the error tokens decrease strikingly along with the increase of the T-I distance. In nearly 80 percent of the errors, a potential intruder can be found within 2 syllables from the target syllable, whereas errors with a T-I distance longer than 5 syllables account for only fewer than 4 percent.

To summarize, the results confirmed the effect of contextual availability in TSM speech errors. Most errors are attributed to the influence of the surrounding segments in the utterance. The domain of the availability effect is limited in 7 syllables from the target. In addition, the error tokens are inversely proportional to the T-I distance. The shorter the distance of two segments involved in errors, the higher the error rate.

## 5. Analysis 2: Phonetic similarity

### 5.1 Measurement of feature differences

Three featural systems are adopted to measure phonetic similarity between consonants and to examine possible differences these featural systems may have on the measurement. The three systems are revised from Jaeger (1992), Wan (1999), and Chomsky & Halle (1968). Jaeger (1992) develops a featural system to explain consonant interactions in English errors in terms of similarity. This system comprises five featural dimensions: ‘place of articulation’, ‘continuancy’, ‘frication’, ‘voice’, and ‘nasality’. The feature ‘place of articulation’ is a multi-value feature, while all the others are binary features. Based on Jaeger (ibid), Wan (1999) uses a revised five featural system to deal with speech errors in Taiwan Mandarin. She replaces the binary feature ‘voice’ with a trinary feature ‘voice onset time’ to discriminate aspirated, unaspirated and voiced consonants. As mentioned in Section 2, nasal consonants are allophones of voiced plosives in TSM. Hence, in this study, we adopt a revised five-featural system from Jaeger (ibid) by replacing the feature ‘nasality’ with ‘aspiration’, and a revised four-featural system from Wan (ibid) by discarding the feature ‘nasality’. In addition, we also adopt a six-featural system from Chomsky & Halle (1968) by representing the place of articulation with two binary features ‘anterior’ and ‘coronal’ for comparison. The three featural systems are abbreviated as FD(4), FD(5), and FD(6) according to the number of their feature components. The feature values of TSM consonants based on the three featural systems are given in Tables 4-6.

As mentioned previously, phonetic similarity of two segments is measured indirectly via counting feature differences shared by the two segments. Take the two segments [b] and [l] in Table 4 as an example to illustrate the measurement of feature differences. The two segments only differ in the feature ‘Place’. Therefore the number of feature differences in this consonant pair [b-l] is counted as one.

Feature differences shared by the target and intruding segments (hereafter the T-I pair) interacting in speech errors are measured using the above-mentioned three featural systems. The results of error tokens sorted by the number of feature differences are given in Tables 7-9, based on FD(4), FD(5), and FD(6), respectively.

**Table 4. Feature values of TSM consonants based on FD(4)***

| consonant | **b** | **l** | **g** | **p** | **t** | **k** | **p$^h$** | **t$^h$** | **k$^h$** | **ts** | **ts$^h$** | **dz** | **s** | **h** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Place | L | D | V | L | D | V | L | D | V | D | D | D | D | G |
| VOT | V | V | V | U | U | U | A | A | A | U | A | V | U | U |
| Fri | - | - | - | - | - | - | - | - | - | + | + | + | + | + |
| Cont | - | - | - | - | - | - | - | - | - | - | - | - | + | + |

*Abbreviations in Tables 4-6, Place: Place of articulation, containing four categorical values: ‘labial’ (L), ‘dental’ (D), ‘velar’ (V), and ‘glottal’ (G); VOT: Voice Onset Time, composed of three values, ‘voiced’ (V), ‘unaspirated’ (U), and ‘aspirated’ (A); Fri: Friction; Cont: Continuancy; Voi: Voice; Asp: Aspiration; Ant: Anterior; Cor: Coronal; Del rel: Delayed release.

**Table 5. Feature values of TSM consonants based on FD(5)**

| consonant | **b** | **l** | **g** | **p** | **t** | **k** | **p$^h$** | **t$^h$** | **k$^h$** | **ts** | **ts$^h$** | **dz** | **s** | **h** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Place | L | D | V | L | D | V | L | D | V | D | D | D | D | G |
| Voi | + | + | + | - | - | - | - | - | - | - | - | + | - | - |
| Asp | - | - | - | - | - | - | + | + | + | - | + | - | - | - |
| Fri | - | - | - | - | - | - | - | - | - | + | + | + | + | + |
| Cont | - | - | - | - | - | - | - | - | - | - | - | - | + | + |

**Table 6. Feature values of TSM consonants based on FD(6)**

| consonant | **b** | **l** | **g** | **p** | **t** | **k** | **p$^h$** | **t$^h$** | **k$^h$** | **ts** | **ts$^h$** | **dz** | **s** | **h** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ant | + | + | - | + | + | - | + | + | - | + | + | + | + | - |
| Cor | - | + | - | - | + | - | - | + | - | + | + | + | + | - |
| Voi | + | + | + | - | - | - | - | - | - | - | - | + | - | - |
| Asp | - | - | - | - | - | - | + | + | + | - | + | - | - | - |
| Del rel | - | - | - | - | - | - | - | - | - | + | + | + | - | - |
| Cont | - | - | - | - | - | - | - | - | - | - | - | - | + | + |

**Table 7. Frequency of error tokens based on FD(4)**

| Feature differences | 1 | 2 | 3 | 4 | Total |
|---|---|---|---|---|---|
| Tokens | 162 | 76 | 19 | 16 | 273 |
| % | 59.34 | 27.84 | 6.96 | 5.86 | 100 |

**Table 8. Frequency of error tokens based on FD(5)**

| Feature differences | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|
| Tokens | 164 | 72 | 21 | 16 | 0 | 273 |
| % | 60.07 | 26.37 | 7.69 | 5.86 | 0 | 100 |

**Table 9. Frequency of error tokens based on FD(6)**

| Feature differences | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| Tokens | 112 | 83 | 63 | 15 | 0 | 0 | 273 |
| % | 41.03 | 30.40 | 23.08 | 5.49 | 0 | 0 | 100 |

The results based on the three featural systems show similar distribution patterns except for the discrepancy in the percentage of error tokens. The number of error tokens is inversely related to the number of feature differences of the T-I pairs. The fewer feature differences two segments in the T-I pairs share, the more the error tokens. T-I pairs with only one feature difference contribute to the most error tokens. Error tokens decrease as the feature differences of T-I pairs increase.

Concerning the percentage of error tokens with different feature differences, the results of FD(4) and FD(5) also show a similar distribution pattern whereas those of FD(6) reveal a large discrepancy in the percentage of error tokens with 1 and with 3 feature differences. Error tokens of the T-I pairs with 1 feature difference are comparatively fewer, and error tokens of the T-I pairs with 3 feature differences are relatively more in FD(6). This discrepancy is mainly attributed to the representation of the place of articulation. Both FD(4) and FD(5) use one feature to represent the place of articulation whereas FD(6) adopts two. Hence, for T-I pairs in which the two segments differ in the place of articulation, some of them are counted as with 1 feature difference in FD(4) and FD(5), but as with 2 feature differences in FD(6), thereby resulting in the lower percentage of T-I pairs with 1 feature difference and the higher percentage of T-I pairs with 2 and 3 feature differences in FD(6).

In the ensuing discussion, we will assess the appropriateness of the three featural systems as an index of phonetic similarity with our error data, and select the best one as an index for a further analysis of TSM errors if there is any.

## 5.2 Assessment of the three featural systems

The significant difference between the featural system FD(6) and the other two FD(5) and FD(4) mainly lies in the representation of the place of articulation. The former adopts two binary features 'anterior' and 'coronal' to represent this property, while the latter two use a multi-value feature 'Place of articulation' to display the same function. Concerning FD(5) and FD(4), the difference between these two systems lies in the discrimination of voicing and aspiration. FD(5) adopts two binary features 'voicing' and 'aspiration', while FD(4) adopts a trinary feature 'VOT' to achieve this function. The degree of phonetic similarity of two consonants may vary according to the features used as an index, thereby resulting in different predictions on error tokens. Hence, crucial data for comparison among the featural systems lie in consonant errors that can discriminate the essential differences among these featural systems.

The results shown in Tables 7-9 above reveal a clear inverse proportion between the feature differences of T-I pairs and the number of error tokens. The fewer the feature differences of T-I pairs, the more the error tokens. It implies that given two T-I pairs, the pair with more feature differences should less frequently involve errors than the other with fewer feature differences do. Error data that can discriminate the fundamental differences among the three featural systems are analyzed to see if the distribution pattern shown in the sub-set of consonant errors supports this implication derived from the general pattern.

First, FD(6) is compared with FD(5) and FD(4) based on three sets of consonants differing by place of articulation, i.e. the voiced stops [b, l, g], the unaspirated stops [p, t, k], and the aspirated stops [$p^h$, $t^h$, $k^h$]. The feature differences shared by any two segments within a set based on each featural system and the error tokens involving these two consonants are listed in Table 10.

**Table 10. Feature differences of consonant pairs based on three featural systems and error tokens involving these consonant pairs**

| Consonant pair | | Voiced stops | | | Unaspirated stops | | | Aspirated stops | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | [b-l] | [b-g] | [l-g] | [p-t] | [p-k] | [t-k] | [$p^h$-$t^h$] | [$p^h$-$k^h$] | [$t^h$-$k^h$] |
| Feature difference | FD(6) | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 |
| | FD(5) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | FD(4) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Error tokens | | 3 | 2 | 5 | 5 | 12 | 25 | 1 | 0 | 7 |

The results in Table 10 show that the labial-dental and the labial-velar pairs in each consonant set, i.e. [b-l] and [b-g] in the voiced stops, [p-t] and [p-k] in unaspirated stops, and [$p^h$-$t^h$] and [$p^h$-$k^h$] in the aspirated stops, all differ in one feature in the three featural systems. The difference among these three featural systems lies in the measurement of the dental-velar pair in each consonant set. The three dental-velar pairs – [l-g], [t-k], and [$t^h$-$k^h$] – share two feature differences based on FD(6), but only one feature difference based on FD(5) and FD(4). That is, the dental-velar consonant pair in each set bears more feature differences than the labial-dental and the labial-velar pairs based on FD(6). According to the distribution pattern of an inversely proportional relation between the feature differences of T-I pairs and the number of error tokens, FD(6) predicts fewer error tokens involving the dental-velar pair than involving the labial-dental or the labial-velar pair in each set. There is no such implication based on FD(5) and FD(4) as all consonant pairs differ by the same number of feature differences.

The error tokens of each consonant pair are given in the last row of Table 10. The dental-velar consonant pair in each set contributes to more error tokens than the labial-dental and the labial-velar counterparts, revealing a direct proportion between feature differences of T-I pairs and error tokens. This distribution pattern shown in these sub-sets of consonant errors is obviously contrary to the prediction of FD(6) and also inconsistent with the pattern revealed in all error data. Accordingly, FD(6) is unable to properly index the phonetic similarity of TSM consonants involved in speech errors concerning the consistency on the distribution patterns.

The same rationale is applied to compare FD(5) and FD(4) by examining the interaction of another three sets of consonants in speech errors: the labial stops [b, p, $p^h$], the dental stops [l, t, $t^h$], and the velar stops [g, k, $k^h$]. The consonants in each set differ in voicing or aspiration or both. The feature differences of any two segments within each set based on FD(5) and FD(4) and the error tokens involving the two segments are listed in Table 11.

**Table 11. Feature differences of consonant pairs based on FD(5) and FD(4) and error tokens involving these consonant pairs**

| **Consonant pair** | | Labial stops | | | Dental stops | | | Velar stops | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | [b-p] | [p-$p^h$] | [b-$p^h$] | [l-t] | [t-$t^h$] | [l-$t^h$] | [g-k] | [k-$k^h$] | [g-$k^h$] |
| Feature differences | FD(5) | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 |
| | FD(4) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Error tokens | | 2 | 1 | 0 | 17 | 5 | 0 | 0 | 19 | 0 |

It is clear in Table 11 that the voiced-aspirated consonant pair in each set, i.e. [b-$p^h$], [l-$t^h$], and [g-$k^h$], bears more feature differences than the voiced-unaspirated and unaspirated-aspirated counterparts based on FD(5), while every pair of consonants bears the same number of feature differences based on FD(4). Hence, FD(5) predicts fewer errors involving the voiced-aspirated consonant pair than the other two pairs in each set according to the distribution pattern shown in all data.

The prediction of FD(5) is supported by the distribution of error tokens shown in the last row of Table 11. The voiced-aspirated consonant pair involves the fewest errors in each set. As all the consonant pairs in each set share the same number of feature differences based on FD(4), there is no such distributional prediction on error tokens in FD(4) in terms of phonetic similarity. Therefore, FD(5) seems a better featural system than FD(4) to account for the phonetic similarity of TSM consonants in speech errors as it properly and consistently represents the general and detailed distribution patterns of error tokens.

In the following analysis, we adopt FD(5) as an index of phonetic similarity to further compare the distribution pattern shown in speech errors and in TSM consonant inventory to see if they bear any distributional correlation.

### 5.3 Comparison of the token distribution in speech errors and in the consonant inventory in terms of phonetic similarity

The general results from the measurement of feature differences among consonants and error tokens reveal that error tokens are inversely proportional to the number of feature differences two segments share. To interpret this distributional pattern in terms of phonetic similarity, the higher the phonetic similarity between two segments, the higher the chance these two segments involve interaction in errors. Yet, before such conclusion is drawn, a further comparison is made between the distributional results of T-I pairs in speech errors and that of the consonant pairs in the consonant inventory to rule out the possibility of frequency effect from the consonant inventory. Hence, feature differences among any two segments in the consonant inventory and the token distribution of consonant pairs with different number of feature differences are also calculated.

The distribution of tokens based on the feature differences of T-I pairs in speech errors, as shown in Table 8, is repeated in Table 12 with one more column of cumulative percentage. The phonetic similarity among consonants in the TSM consonant inventory is also measured for comparison. There are 14 consonants included for analysis in this study.[5] Given any two consonants to form a pair, there are in total 91 consonant pairs. The tokens of consonant pairs with different numbers of feature differences are shown in Table 13. For example, in column 2 of Table 13, there are 22 consonant pairs in the TSM consonant inventory in which two consonants share 1 feature difference. The distribution patterns shown in Table 12 and 13 are compared in order to see if there is any relation between them.

**Table 12. Tokens of errors based on the feature differences of T-I pairs in speech errors**

| Feature differences | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|
| Tokens | 164 | 72 | 21 | 16 | 0 | 273 |
| % | 60.07 | 26.37 | 7.69 | 5.86 | 0 | 100 |
| Cumulative % | 60.07 | 86.44 | 94.13 | 100 | 100 | 100 |

**Table 13. Tokens of consonant pairs based on the feature differences of consonant pairs in the TSM inventory**

| Feature differences | 1 | 2 | 3 | 4 | 5 | Total |
|---|---|---|---|---|---|---|
| Tokens | 22 | 30 | 25 | 14 | 0 | 91 |
| % | 24.18 | 32.97 | 27.47 | 15.39 | 0 | 100 |
| Cumulative % | 24.18 | 57.15 | 84.62 | 100 | 100 | 100 |

[5] As mentioned in Section 2, the glottal stop is not distinctive in the onset position.

A comparison on the token distribution between the above two tables clearly reveals several significant differences. These differences indicate that the token distribution shown in speech errors is not related to or influenced by the token distribution shown in the consonant inventory.

First of all, in Table 12, there is a clear inverse correlation between the error tokens in speech errors and the number of feature differences, whereas there is no such correlation between tokens of consonant pairs in the consonant inventory and the number of feature differences shown in Table 13. The error tokens in speech errors are inversely proportional to the number of feature differences. The error tokens decrease strikingly as the number of feature differences increases. On the contrary, the tokens of consonant pairs in Table 13 are randomly distributed. Consonant pairs with two feature differences are the most frequent in terms of tokens, followed by consonant pairs with three feature differences, which are followed by consonant pairs with 1 and 4 feature differences in turn.

Secondly, in the speech-error data, T-I pairs with one feature difference are the most frequent, and account for 60% of the total tokens. That is, more than half of speech errors involve two segments differing by only one feature. However, in the consonant inventory, consonant pairs sharing one feature difference are less frequent than consonant pairs with two or three feature differences, and only contribute to less than one-fourth of the total tokens.

Thirdly, nearly ninety percent of T-I pairs in speech errors differ in only one or two features. In the consonant inventory, however, consonant pairs differing in one or two features only account for slightly more than half of the total consonant pairs (57%). The average number of feature differences shared by two segments involving speech errors is 1.59 feature differences, much smaller than that shared by two segments in the consonant inventory, which is 2.34 feature differences.

Lastly, although T-I pairs in the speech errors and consonant pairs in the consonant inventory both differ in four features at most, there is a great discrepancy in the percentage of tokens with four feature differences in these two sets of data. There are less than six percent of T-I pairs in speech errors differing by four features, while there are more than fifteen percent of consonant pairs in the consonant inventory differing by four features.

In summary, the four distributional differences of tokens between the speech errors and the consonant inventory mentioned above rule out the frequency effect or influence of the latter on the former. Furthermore, a linear correlation test also shows statistical significance on the inverse correlation between the feature differences and the error tokens ($p< .05$). Accordingly, in terms of phonetic similarity, the results support the effect of phonetic similarity on the occurrence of speech errors. The error tokens are directly proportional to the degree of phonetic similarity shared by two segments involving errors.

**6. Discussion**

Concerning the availability effect, the results shown in TSM speech errors are consistent with those found in different languages, both intonation languages and tone languages, in the distribution pattern of error tokens and the domain of contextual influence (Cohen 1973, Nooteboom 1973, Wan 2007). Most of the consonant errors are attributed to the influence of surrounding consonants in the utterance. Also, the error tokens are inversely proportional to the increase of the T-I distance, and the intruding segments are located within a limited domain of 7 syllables from the target. The cross-language finding on the limited domain of contextual influence has a theoretically significant implication for the processing of speech production.

There are various proposals delineating the units involved in the process of speech production such as information block, message, sentence, phrase or tone group (Fodor, Bever, & Garrett 1974, Garrett 1980, Bock 1982, Levelt 1989). Levelt (1989) argues that there is no single unit preserved intact thoroughly in the processing of speech production. Different processing components involve different processing units. For instance, the processing stage of grammatical encoding might involve units such as 'noun phrase' and 'verb phrase'. These elements, nevertheless, will be decomposed and reconstructed to different units in later stages of processing such as phonological encoding. The domain of the contextual influence on phonological errors, i.e. 7 syllables, corresponds to the limited domain of the working memory span, which is proposed to contain about 7 units (Miller 1962). The finding regarding the domain of the contextual influence within 7 syllables from the target syllable probably also provides an argument for syllables as a basic unit in the processing stage of phonemic programming (Cohen 1973, Nooteboom 1973, Levelt 1989). In addition, this consistent finding shown in both intonation and tone languages also further implies syllables as a common basic unit in some processing stage of speech production in these typologically different languages.

Besides the common findings, there is also a discrepancy in TSM and previous literature. The percentage of contextual errors shown in our study is higher than that shown in English. It is also the same in Wan's (2007) study on speech errors of Taiwan Mandarin. This cross-language discrepancy might be due to the different phonological structures between English and Chinese languages. We attribute the higher percentage of contextual errors in TSM and Taiwan Mandarin to the relatively smaller amount of consonants in these two languages. Given two consonants X and Y, there are fewer consonants in TSM than in English. Therefore, there is a higher chance for the two consonants to interact in speech errors in TSM than in English for they have a higher frequency of occurrence in an utterance. Also, there is a higher chance of finding more than one Y as a potential intruder in the utterance of TSM than of English for the same reason. This explanation is partially supported by the finding in this study that over one-third of the errors contain more than one potential intruder within 7 syllables from the target in an utterance. Yet, further confirmation with the average number of potential intruder in English speech errors is needed.

Next, the comparison among the three featural systems used as an index of the phonetic similarity shows that the adoption of different systems does lead to different results. Nevertheless, the difference does not lie in the general tendency of token distribution but in the consistency of the tendency shown in the error data. FD(6) adopts two features ‘anterior’ and ‘coronal’ to represent the place of articulation of consonants. Examination of voiced, unaspirated, and aspirated plosives with different place of articulation shows that FD(6) reveals contrary distribution patterns between these sub-set data and the total error data in terms of the effect of phonetic similarity. Concerning FD(5) and FD(4), the former adopts two features ‘voicing’ and ‘aspiration’ to discriminate articulatory manners of consonants, while the latter combines the two features as a multi-valued feature ‘VOT’. Examination of labial, alveolar, and velar plosives shows that FD(5) represents the distribution pattern shown in the sub-set of error data better than FD(4). These discrepancies in these featural systems would not be uncovered without a detailed examination of certain error data that can discriminate the essential differences among these featural systems.

Yet, this issue has been much overlooked in the study of speech errors. Most studies on phonetic similarity are based on a featural system selected without comparison (Jaeger 1992, Wan 1999, Liu 2009). This selected featural system may represent consonants in the target language well, and the analysis of phonetic similarity based on this featural system also represents a correlation between the phonetic similarity and error tokens as well, just like FD(6) and FD(4) do in this study. Nevertheless, the comparison among three featural systems in this study shows that the adoption of a featural system as an index of phonetic similarity without comparison in advance might lead to some discrepant or even controversial results hidden in the sub-set data, and these critical hidden results are usually overlooked.

Moreover, the comparison among the featural systems also discloses a fundamental problem concerning the representation of a feature value. Take the three consonants [ts], [$ts^h$], and [s] in TSM for example. Phonologically, [ts] and [$ts^h$] are grouped in a category ‘affricate’, and [s] is labelled as another category ‘fricative’. Affricate contrasts with fricative by the feature ‘continuancy’. Both [ts] and [$ts^h$] are [-

continuant], and [s] is [+continuant]. The two affricates are further discriminated by the feature ‘aspiration’. The consonant [ts] is [-aspirated], while [$ts^h$] is [+aspirated]. The feature ‘aspiration’ is irrelevant to the phonetic property of fricatives. However, according to the conventional feature theory with full specification, a consonant should be labelled with a value in all featural dimensions, either [+] or [-]. The default value is [-] if a consonant does not carry the property of a given feature. Therefore, the consonant [s] is labelled with a [-] value, the same as [ts], in the featural dimension ‘aspiration’.

Based on the conventional representation of the feature value, the unaspirated affricate [ts] and the fricative [s] differ by one feature, i.e., ‘continuancy’. The aspirated affricate [$ts^h$] and the fricative [s] differ by two, i.e., ‘continuancy’ and ‘aspiration’. Thereby [ts] and [s] have a higher degree of phonetic similarity than [$ts^h$] and [s] do, according to the featural system. However, phonetically speaking, the aspirated affricate [$ts^h$] and the fricative [s] seem more similar than the unaspirated affricate [ts] and [s] as the former pair cause a high-frequency noise in perception. This phonetic property is also reflected in our error data. The fricative [s] is more frequently replaced by [$ts^h$] than by [ts]. In other words, the aspiration property actually makes an aspirated affricate phonetically more similar to a fricative than an unaspirated affricate. But in the phonological featural system, the feature ‘aspiration’ turns out to make an aspirated affricate phonetically less similar to a fricative than an unaspirated affricate. The inconsistency between the error data and the measurement of phonetic similarity based on the featural system implies a need of reconsideration regarding the default setting of a feature value. The number of error tokens in our corpus so far is not enough for a detailed discussion on this issue. Hence, we merely point out the phenomenon here to raise the awareness. It is an issue worthy of further research.

## 7. Conclusion

In this study, we explore the effects of contextual availability and phonetic similarity on the consonants errors of TSM. The results show supportive findings to both factors.

Concerning the effect of contextual availability, the error tokens are inversely proportional to the T-I distance. Error tokens decrease strikingly along with the increase of the T-I distance. Moreover, the domain of contextual influence is limited within 7 syllables from the target, consistent with the results shown in previous literature. The distance of 7 syllables shown in cross-language studies is of theoretical significance concerning the linguistic unit in language processing. It provides a supportive argument for syllables as a basic linguistic unit in language processing.

Regarding the effect of phonetic similarity, the results show a direct correlation between the degree of phonetic similarity shared by two segments involved in speech erorrs and the number of error tokens. Two segments sharing more phonetic similarity tend to involve interaction in speech errors more often. The results are consistent with previous studies on different languages, implying that these effects probably have a universal tendency in terms of speech errors. In addition, the comparison among different featural systems shows that the effect of phonetic similarity might vary depending on the featural system selected as an index of similarity, an issue which has been scarcely discussed. It indicates that researchers should be more cautious in the adoption of a feature system as an index in linguistic study.

## References


Berg, Thomas. 1987. *A Cross-Linguistic Comparison of Slips of the Tongue*. Bloomington: Indiana University Linguistics Club.

Bock, J. Kathryn. 1982. Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review* 89.1:1-47.

Boomer, Donald S., and John D. M. Laver. 1968. Slips of the tongue. *International Journal of Language & Communication Disorders* 3.1:2-12.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Cohen, Anthony. 1973. Errors of speech and their implication for understanding the strategy of language users. *Speech Errors as Linguistic Evidence*, ed. by Victoria Fromkin, 88-92. The Hague: Mouton.

Cutler, Anne. (ed.) 1982. *Slips of the Tongue and Language Production*. New York: Mouton.

Dell, Gary S., and Peter A. Reich. 1977. A model of slips of the tongue. Paper presented at The Third LACUS Forum, 448-455. Columbia, SC: Hornbeam Press.

Fay, David, and Anne Cutler. 1977. Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry* 8.3:505-520.

Fodor, Jerry A., Thomas G. Bever, and Merrill F. Garrett. 1974. *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw-Hill.

Fromkin, Victoria. (ed.) 1973. *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

Fromkin, Victoria. 1989. Grammatical aspects of speech errors. *Linguistics: The Cambridge Survey,* vol. 2: *Linguistic Theory: Extensions and Implications*, ed. by Frederick J. Newmeyer, 117-138. Cambridge: Cambridge University Press.

Garrett, M. F. 1980. Levels of processing in sentence production. *Language Production* 1:177-220.

Jaeger, Jeri J. 1992. Phonetic features in young children's slips of the tongue. *Language and Speech* 35.1-2:189-205.

Jaeger, Jeri J. 2005. *Kid's Slips: What Young Children's Slips of the Tongue Reveal about Language Development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Levelt, W. J .M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: The MIT Press.

Liu, Joyce H.-C. 2009. Speech Errors in Taiwan Southern Min: A Phonological Study. Doctoral dissertation, National Tsing Hua University, Taiwan.

MacKay, Donald G. 1970. Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia* 8.3:323-350.

Miller, George A. 1962. Decision units in the perception of speech. *Institute of Radio Engineers Transactions on Information Theory* 8.2:81-83.

Nooteboom, Sieb G. 1973. The tongue slips into patterns. *Speech Errors as Linguistic Evidence,* ed. by Victoria Fromkin, 144-156. The Hague: Mouton.

Shattuck-Hufnagel, Stefanie. 1985. Context similarity constraints on segmental speech errors: An experimental investigation of the role of word position and lexical stress. *Proceedings of the Conference on the Planning and Production of Speech in Normal and Hearing-impaired Individual: A Seminar in Honor of S. Richard Silverman*, ed. by Judith L. Lauter, ASHA report 15:43-49. Maryland: American Speech-Language-Hearing Association (MSHA).

Shattuck-Hufnagel, Stefanie. 1986. The representation of phonological information during speech production planning: Evidence from vowel errors in spontaneous speech. *Phonology Yearbook* 3:117-149.

Shattuck-Hufnagel, S., and D. H. Klatt. 1979. The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior* 18.1:41-55.

Stemberger, Joseph Paul. 1989. Speech errors in early child language production. *Journal of Memory and Language* 28.2:164-188.

Sturtevant, Edgar Howard. 1947. *An Introduction to Linguistic Science*. New Haven: Yale University Press.

Van den Broecke, M. P. R., and L. Goldstein. 1980. Consonant features in speech errors. *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand,* ed. by Victoria Fromkin, 47-65. New York: Academic Press.

Wan, I-Ping. 1999. Mandarin Phonology: Evidence from Speech Errors. Doctoral dissertation, State University of New York at Buffalo, U.S.A.

Wan, I-Ping. 2007. Mandarin speech errors into phonological patterns. *Journal of Chinese Linguistics* 35.2:185-224.

Yang, Xiu-Fang. 1995. *Min Nan Hua Shou Gao* [Manuscripts of the Grammar of Taiwan Southern Min]. Taipei: Da-an Press.

[Received June 8, 2012; revised September 6, 2012; accepted Septermber 11, 2012]

Department of Digital Applied Chinese
Kai Nan University
Taoyuan, TAIWAN
Joyce H.-C. Liu: jliu0227@gmail.com

# 語境可用性以及語音相似度在語誤當中的影響

劉慧娟
開南大學

本文探討語境可用性以及語音相似度兩個因素對台閩語子音音段語誤的影響。結果顯示大部分的語誤都是受到語境當中的的語音所影響，而且語誤的數量與目標音跟來源音之間的距離成反比關係；語境可用性範圍侷限在目標音七個音節內的距離。此現象對於探討大腦處理語言訊息時所運用的語言單位之相關研究具有顯著意義。語音相似度分析顯示語誤數量與目標音跟來源音兩者的語音相似度成正比關係，兩個語音的語音相似度越高，越容易發生語誤。總言之，本文的分析結果與文獻上其他語言之語誤研究結果一致，顯示語誤有跨語言的普遍性特徵；同時，本文也進一步發現語音相似度的影響會隨研究所採取的不同語音徵性系統而有程度上的差別。

關鍵詞：語誤、子音、語境可用性、語音相似度、台閩語