I-PING WAN AND JERI JAEGER

THE PHONOLOGICAL REPRESENTATION OF TAIWAN MANDARIN VOWELS: A PSYCHOLINGUISTIC STUDY

One of the fundamental goals of every phonological theory is to account for the nature of the basic units of speech sounds, and the relationships between these units and their contextual variants. This relationship is equally crucial to phonological theory whether it is called 'phonemes and allophones', 'underlying and surface forms', or 'input and output'. However, purely structural analyses of phonological systems can often produce several hypotheses regarding the phonemic inventory and its surface reflexes in any particular language, all of which are supportable by the contrast and alternation patterns of the language. In this paper we look at four such hypotheses regarding the underlying vowel system of Mandarin, all based on Beijing Mandarin: the six-vowel system of C. Cheng (1973), the five-vowel systems of R. Cheng (1966) and of Lin (1989), and the four-vowel system of Wu (1994). We then present distributional, phonetic, and psycholinguistic evidence (the latter based on a corpus of 238 syntagmatic speech errors or 'slips of the tongue' involving vowels) that the vowel system of the dialect of Mandarin currently spoken in Taiwan cannot be accounted for by any of these hypotheses. We then propose a new 5-vowel system for Taiwan Mandarin, based on the distributional, phonetic, and especially the psycholinguistic facts. We conclude that phonological theories which are compatible with psycholinguistic evidence such as the data presented here are those most likely to be modeling the actual cognitive representations and processes of real speakers.

1. Introduction

One of the fundamental goals of any phonological theory is to be able to account for the nature of the basic units of speech sounds, and the relationship between these units and their contextual variants, both universally and in particular languages. Whether we talk about this relationship in terms of 'phonemes and allophones', or 'underlying and surface forms', or 'input and output', we are making the same basic assumption: In any language there is an inventory of possible abstract sound categories in which words are represented, which will be manifested differently in different phonetic contexts due to both universal and language-specific factors (Goldsmith (1995: 2)). In psycholinguistics we would talk about these same phenomena in terms of the units in which phonological form is represented in lexical entries, and the relationships between these units to either the incoming acoustic signal, or to their manifestation in contextual speech. In any particular language, the actual analysis of the abstract phonological units and their contextual manifestations is frequently the source of much theoretical controversy.

This is certainly true of the system of vowels in Mandarin Chinese. There

Journal of East Asian Linguistics **12**, 205–257, 2003.

has been a longstanding controversy in the literature over the number of underlying vowel categories in Mandarin, and the relationship of the myriad of surface vowel forms to these phonemic categories (e.g., Chao (1934, 1968), R. Cheng (1966), C. Cheng (1973), Pullyblank (1983), Lin (1989), Wang (1993), Wu (1994)). The reason for the continuance of this controversy is that most phonetic manifestations of vowels in Mandarin occur in a fairly narrow range of contexts, which suggests that they probably can be reduced to a smaller set of basic vowel categories. However, a standard structural analysis of the contexts in which the various vowel alternates occur allows for several different resolutions to the problem, each one of which can be motivated on the basis of theoretical and language-internal consistency. Thus if one uses strictly 'internal' evidence (Ohala (1986)), it may be the case that this controversy cannot be definitively resolved. Furthermore, while many of these theories regarding the underlying vowel system intend to capture the basic vowel system for all dialects of Mandarin, they are for the most part based on the Beijing dialect. Thus even if the analysis adequately captures that dialect, it may be the case that other Mandarin dialects are different enough from Beijing Mandarin that they can be accounted for only by a different underlying vowel system.

The purpose of the present paper is to bring some distributional, phonetic, and particularly psycholinguistic evidence to bear on this issue, looking specifically at the vowel system of the dialect of Mandarin spoken in Taipei, Taiwan (hereafter TM). First we will review four theories regarding the underlying vowel system of Mandarin, which are intended to be applicable to all dialects of Mandarin: the 6 vowel system of C. Cheng (1973), the five-vowel systems of R. Cheng (1966) and Lin (1989, 1992), and the four-vowel system of Wu (1994). We then look at some distributional and phonetic facts regarding the vowel system of TM, and show that certain sound changes have caused the Taiwan and Beijing dialects to have different vowel systems, particularly in terms of the high central unrounded (or 'apical') vowels, and in terms of the phonemic affiliation of the lower-mid front unrounded vowel [ɛ]. We then argue for this new hypothesis regarding the underlying system of TM by presenting evidence from a corpus of 238 syntagmatic phonological speech errors or 'slips of the tongue', where either one vowel is substituted for another, or where the environment adjacent to a vowel is changed, causing the vowel to shift in quality. The questions we ask are the following:

(1) Which vowels can be substituted for each other in syntagmatic phonological errors? The rationale for this question is that if two vowel phones can be substituted for each other without any change in the contextual

environment, then they must derive from different underlying categories, since slips of the tongue never produce illegal surface phonetic forms.

(2) How do vowel phones change if the contiguous environment is changed by an error? The rationale for this question is that if the intended vowel phone is changed to a different vowel phone consistently when the context changes in some particular way, then these two phones are likely to be contextually determined manifestations of the same abstract vowel category.

This paper is organized as follows. In the following section we will discuss the four relevant theories mentioned above regarding the basic vowel categories of Mandarin, and then lay out the facts regarding the vowel phones in TM, including the restrictions on their contextual occurrences and ways in which they differ from Beijing Mandarin. In the third section we will present our methodology for the collection and analysis of slips of the tongue in detail. Section four presents our psycholinguistic results, and the analysis of those results both in terms of the four competing hypotheses and in terms of our proposed model for TM. In section five we summarize our study, and discuss the phonological analysis supported by this study in detail, relating it to a psycholinguistic model of speech production planning.

Because we want to make it clear that the issues being addressed in this paper are relevant to a phonological analysis of Mandarin regardless of the phonological theory in which such an analysis is performed, we will be eclectic in our terminology throughout this paper. That is, we will use the terms ‘phoneme’, ‘underlying representation’, ‘input’, ‘abstract phonemic category’, and ‘cognitive category’ interchangeably to indicate the abstract form in which vowels are represented in lexical entries, and we will use the terminology ‘allophone’, ‘surface form’, and ‘phonetic manifestation’ interchangeably to indicate the form of the vowel which is actually spoken or is the ‘output’ of the phonological ‘rules’ or ‘constraints’ component during processing. We take the stance that, from a psycholinguistic point of view, all of these ways of talking about vowels should be reducible to the same cognitive entities.

2. Mandarin Vowels

2.1. *Previous Theoretical Analyses*

Out of the many analyses of the Mandarin vowel system to appear previously, we have selected the following four proposals to review here: C.

Cheng (1973), who proposed a six-vowel system based on phonemic evidence; R. Cheng (1966) who proposed a five vowel system also based on phonemic analysis; Lin (1989), whose 5-vowel system is developed from autosegmental phonology and underspecification theory; and Wu (1994), whose 4-vowel system is based on a feature geometry analysis. We selected these four theories for three reasons. First, they cover the range of hypotheses regarding the number of underlying vowels in Mandarin, from 4–6.[1] Second, they represent a range of types of phonological theories, from phonemic analyses through autosegmental and feature geometry approaches. Third, they all propose to represent the basic structure of the Mandarin vowel system regardless of dialect, and thus it is appropriate for us to compare the vowel system of TM to these proposals.

C. Cheng's (1973) analysis begins with the assumption that there are 12 surface vowels in Mandarin: [i, ɿ, ʅ, y, u, e, ɛ, ə, o, ɤ, a, ɑ]. (All of these symbols are used according to their IPA value, except the two 'apical' vowels: [ɿ], which is a high central unrounded vowel, sometimes called a 'dental' vowel, and [ʅ], a high central retroflexed vowel.) His 6-vowel phonemic system groups allophones as follows:

(1) The 6-vowel system of C. Cheng (1973)
/i/ → [i]
/ɨ/ → [ɿ], [ʅ]
/y/ → [y]
/u/ → [u]
/ɤ/ → [e], [ə], [o], [ɤ]
/ɑ/ → [a], [ɑ], [ɛ]

C. Cheng argues that the ability of a vowel to occur singly or in CV syllables is evidence that the vowel should be analyzed as an underlying phoneme. Since the vowels [i, ɿ/ʅ, y, u, ɤ, a] can occur alone or finally, he treats them as the underlying phonemic centers, as opposed to the other vowel phones which either cannot occur alone or must be conditioned by a preceding or following segment (the details of this conditioning will be discussed in Section 2.2 below). One phonetic difference between his analysis and the other analyses to be discussed here is that he considers the manifestation of the low vowel which occurs in open syllables to be the back vowel [ɑ], whereas all the other theories argue that the low vowel which occurs in open syllables is the central [a]. A phonemic difference is that he considers the two apical vowels as being derived from an underlying high central unrounded vowel phoneme, /ɨ/; it will be shown below that all four theories differ as to the status of the apical vowels. Finally, he considers Mandarin to have vowel sequences rather than vowel-glide or glide-vowel sequences, whereas many other analyses claim that when a high vowel is found adjacent to a non-high vowel, the high vowels are realized as glides.

The 5-vowel system of R. Cheng (1966) assumes the same 12 surface vowels as in the 6-vowel system of C. Cheng (1973); R. Cheng also does not include glides in his system. However, he has one less phoneme due to deriving both [i] and the two apical vowels from the phoneme /i/ as shown in (2).

(2) The 5-vowel system of R. Cheng (1966)
/i/ → [i], [ɿ], [ʅ]
/y/ → [y]
/u/ → [u]
/ə/ → [e], [ə], [o], [ɤ]
/a/ → [a], [ɑ], [ɛ]

Lin (1989) presents a 5-vowel system which is similar to R. Cheng's in that it posits the same 5 phonemes /i, y, ə, a, u/. She makes an argument for selecting these phones as the basic allophones because she claims that they are the phonetic centers of each of the phonemes. However, her analysis differs from R. Cheng's in several ways. First, she states that there are 13 surface vowels, which include all of those in R. Cheng plus the lower-mid back rounded vowel [ɔ]: [i, ɿ, ʅ, y, u, e, ɛ, ə, ɔ, o, ɤ, a, ɑ]. Second, she derives the two apical vowels by rule from an epenthesis and assimilation process; she considers them to be a vocalic extension of the onset consonant, following Chao's (1934) proposal. Third, she considers [ɛ] to be a manifestation of two different phonemes; in the phonetic sequences [jɛn] and [ɥɛn] she argues that the [ɛ] is underlyingly /a/, whereas in the sequences [jɛ] and [ɥɛ], it is underlyingly /ə/. Finally, she does derive glides from high vowels adjacent to non-high vowels, so that, for example, the word /kuai55/ 'obedient' would occur as [kwaj55] in its surface form; similarly the glide [ɥ] is derived from the vowel /y/. Thus her phonemic analysis is as follows.

(3) The 5-vowel system of Lin (1989)
/i/ → [i], [j]
/y/ → [y], [ɥ]
/u/ → [u], [w]
/ə/ → [e/ɛ], [ə], [o/ɔ], [ɤ]
/a/ → [a], [ɑ], [ɛ]
[ɿ, ʅ derived by epenthesis and assimilation rules]

The surface vowels [ɛ] and [ɔ] associated with the /ə/ phoneme are argued to be lowered variants of [e] and [o] respectively. Lin cites as evidence that the [ɛ] in sequences of [palatal glide + ɛ + n] is derived from the /a/ phoneme, the fact that these same morphemes, when they undergo [r] suffixation, surface with the vowel [a] ((1989: 68); see also C. Cheng (1973)). So for example, /ian/ has as its surface form [jɛn]; but for the sequence /ian/ + /r/, the nasal is deleted and the phonetic form is [jar]. The other manifestations of the /ə/ phoneme are derived by assimilation, as will be discussed in detail in Section 2.2 below; Lin's feature system will be discussed in Section 5.3 below.

Finally, Wu (1994) presents a somewhat different analysis of the vowel system of Mandarin, positing four underlying vowels. As in Lin's analysis, Wu derives glides from high vowels, and posits that there are 13 surface vowels; however the set of vowels Wu is working with is somewhat different from the others: [i, ɨ, ɯ, y, u, ə, e, ɤ, o, ɒ, æ, a, ɑ]. Her phonemic groupings are as follows:

(4) The 4-vowel system of Wu (1994)

Simple Vowels	Complex Vowels
/i/ → [i], [ɨ], [ɯ], [j]	/ui/ → [y], [ɥ]
/u/ → [u], [w]	
/ə/ → [e], [ə], [o], [ɒ], [ɤ]	
/a/ → [a], [ɑ], [æ]	

Wu distinguishes the same two apical vowels as in the other theories, but uses the symbols [ɨ] for the dental vowel and [ɯ] for the retroflex vowel. She derives these two vowels from the phoneme /i/ as does R. Cheng, as opposed to setting them up as a separate phoneme (C. Cheng) or deriving them by rule (Lin). For the high front rounded vowel [y], she derives this vowel from a sequence of the two underlying phonemes /ui/. She argues that this vowel should not have the same status as the other high vowels since it has a limited distribution, failing to occur in post-vocalic position (i.e., as a postnuclear glide), where the other high vowels do occur. She considers the lower back rounded vowel which is part of the schwa phoneme to be [ɒ] instead of [ɔ] (e.g., [kwɒ35] 'country'), and analyzes the front raised version of the phoneme /a/ as phonetically [æ] rather than [ɛ] as in the other three theories (e.g., [jæn35] 'language').

This brief review of four diverse hypotheses regarding the basic vowel system of Mandarin shows several areas of difference. Looking first at the high vowels, C. Cheng posits four distinct phonemes, /i, y, ɨ, u/, whereas R. Cheng and Lin posit three phonemes /i, y, u/; R Cheng derives the two apical vowels from /i/, but Lin derives them by rules of epenthesis and assimilation. Wu posits only two high vowels /i, u/; she derives the apical vowels from /i/, but derives [y] from a sequence of /ui/. Thus these analyses raise questions regarding the relationship between the apical vowels and /i/, as well as the status of /y/ as an independent phoneme.

As far as the low vowel is concerned, all four theories consider [a, ɑ] to be conditioned variants of each other; all but C. Cheng consider [a] to be the basic variant, which occurs in open syllables with no conditioning environment. All four theories indicate that there is a front allophone in this set; C. Cheng, R. Cheng, and Lin analyze this segment as [ɛ], whereas Wu calls it [æ]. Lin further argues that there is also an [ɛ] which is derived from the /ə/ phoneme and occurs in open syllables after palatal glides. So for the low vowels, there is some question about which is the central allophone, and about the status of [ɛ].

Finally, with the mid vowels, all four theories consider [e, ə, o, ɤ] to be conditioned variants of this phoneme. The low back rounded vowel [ɔ] does not occur in the theories of C. Cheng or R. Cheng, whereas Wu includes this vowel, saying that it is phonetically [ɒ]. Lin argues that there are lowered versions of [e] and [o] which occur as part of this phoneme as [ɛ] and [ɔ]. Thus the number of surface variants of the mid vowel, as well as the actual phonetic quality of some of these variants, is at issue. The status of [ɛ] is also relevant to this phoneme.

2.2. *Description of Vowel Occurrence and Phonetics in Taiwanese Mandarin*

2.2.1. *Overview*

In the dialect of Mandarin under study here, there are the following 12 surface vowels.

TABLE 1
Vowel phones in Taiwan Mandarin

	Front		Central	Back	
	Unround	Round	Unround	Unround	Round
High	i	y	ɨ		u
Mid	e		ə	ɤ	o
Lower-Mid	ɛ				ɔ
Low			a	ɑ	

The most obvious difference between this set of vowels and the set discussed in Section 2.1 above is that this dialect does not distinguish the two apical vowels (i.e., dental apical [ɩ] vs. retroflex apical [ʅ]). We will return to a detailed discussion of this point below.

The contextual occurrences of these vowel phones are given in Table 2.[2]

As can be seen from Table 2, in TM the five vowels [i, y, u, ɤ, a] can occur alone as complete syllables, and the greatest number of contrasts which can occur with the same onset is also these five, e.g., [li, ly, lu, lɤ, la]. The vowels [i, ɨ, y, u, ɛ, ɔ, ɤ, a] can occur in open syllables; the low vowel in open syllables is [a] rather than [ɑ] in TM. The vowels [i, ɨ, y, ɛ, ə, a] can occur in syllables closed with the nasal [n]; the vowels [i, o, ɑ] occur in syllables closed with the nasal [ŋ]. The vowels [o, ɑ] can occur in syllables closed with the glide [w]; the vowels [a, e] can occur in syllables closed with the glide [j]. A syllable can begin with one of the following vowels: [i, y, ə, o, a, ɑ].

As discussed above, in many traditional and recent analyses of glides, it is claimed that the occurrence of glides can be predicted in every environment by rule (e.g., Lin (1989), Duanmu (1990, 2000), Li (1999)). This is equally true of TM as of other dialects of Mandarin. Glides are analyzed as coming from underlying high vowel phonemes: a high vocalic segment alternates with a glide when adjacent to a nonhigh nucleus vowel, causing an alternation of [i] with [j], [u] with [w], and [y] with [ɥ]; the high central phone [ɨ] does not have a glide alternant.

We will now turn to a detailed discussion of the distribution and phonetics of sets of vowels, organized by vowel height.

2.2.2. *High Vowels*

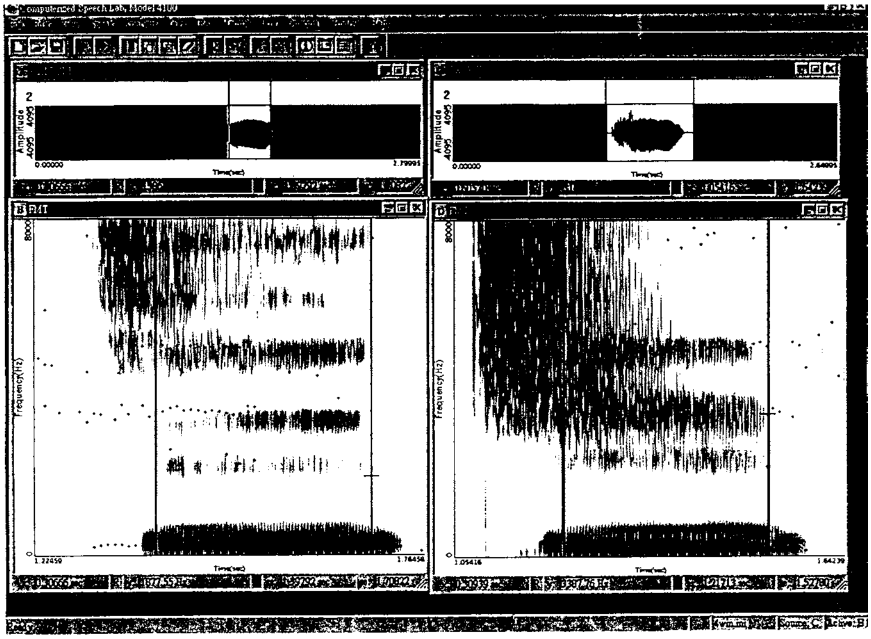
In all of the analyses of Mandarin presented in Section 2.1 above, the researchers have made a distinction between the dental apical vowel [ɩ] or [ɨ] and the retroflex apical vowel [ʅ] or [ɯ]. The major claim has been that the high central vowel after dental obstruent onsets differs in quality from the high central vowel following retroflex obstruent onsets. However, in TM many speakers do not produce strongly retroflexed consonants (see note 6), and more importantly, they do not produce the strongly retroflexed vowel after retroflex consonants which is more typical of the Beijing dialect. Thus these two vowels have fallen together in TM, and are phonetically indistinguishable. The following spectrogram compares the two syllables [tsɨ] and [tʂɨ], in order to illustrate this claim. The utterances in this spectrogram were spoken by a 21-year-old male who is a native speaker of

TABLE 2
Possible co-occurrence of Mandarin vowels with all onsets and codas

	i	ɨ	y	u	e	ɛ	ə	ɔ	o	ɤ	a	ɑ
single	i		y	u						ɤ	a	
_(C/G)	in iŋ		yn				ən		oŋ ow		an aj	ɑŋ ɑw
p, p^{h}	pi pin piŋ			pu	pej		pən		poŋ pow		pa pan paj	pɑŋ pɑw
t, t^{h}	ti tin tiŋ			tu	tej		tən		toŋ tow	tɤ	ta tan taj	tɑŋ tɑw
k, k^{h}				ku	kej		kən		koŋ kow	kɤ	k^{h}a k^{h}an k^{h}aj	kɑŋ kɑw
f				fu	fej		fən		foŋ fow		fa fan	fɑŋ fɑw
s, ts, ts^{h}		sɨ		su	sej		sən		soŋ sow	sɤ	sa san saj	sɑŋ sɑw
ʂ, tʂ, tʂ^{h}, ʐ		ʂɨ		ʂu	ʂej		ʂən		ʂoŋ ʂow	ʂɤ	ʂa ʂan ʂaj	ʂɑŋ ʂɑw
ɕ, tɕ, tɕ^{h}	ɕi ɕin ɕiŋ		ɕy ɕyn									
x				xu	xej		xən		xoŋ xow	xɤ	xa xan xaj	xɑŋ xɑw
m	mi min miŋ			mu	mej		mən		moŋ mow		ma man maj	mɑŋ mɑw
n	ni nin niŋ		ny	nu	nej		nən		noŋ now	nɤ	na nan nɑŋ	nɑŋ nɑw
l	li lin liŋ		ly	lu	lej		lən		loŋ low	lɤ	la lan laj	lɑŋ lɑw
(C)j_						jɛ jɛn			joŋ jow		ja	jɑŋ jɑw
_(C)ɥ_						ɥɛ ɥɛn			ɥoŋ			
(C)w_					wej		wən	wɔ	woŋ		wa wan waj	wɑŋ

TM. (Naturally these spectrograms, as well as those given in (8) below, are meant to serve only as illustrations of the neutralization of this contrast in the dialect of some speakers of TM; full documentation that this neutralization is widespread in TM would of course require data from multiple speakers, with detailed acoustic and statistical analyses.)

(5) Spectrogram comparing [tsɨ] and [tʂɨ]



An analysis of the formant values in the spectrogram showed that that vowel following the dental onset has formant values of approximately F1: 300, F2: 2245, and F3: 3295. For the vowel following the retroflex onset the values are approximately F1: 300, F2: 2285, and F3: 3400. The first formants of the two vowels are nearly identical, and the second formants differ by 80 Hz. The third formant is about 105 Hz higher for the vowel following the retroflex consonant, showing that there is clearly no retroflexion on the vowel, since retroflexion typically causes lowering of all three formants, especially F3 (Jakobson, Fant and Halle (1951), Ladefoged (2001: 212–214)). Since a variation of 100 Hz is within the normal frequency variance for a formant of a single vowel, we would argue that there is no phonetic distinction between the high central vowel after dental vs. retroflex onsets in TM, and thus we notate this vowel as [ɨ] in every case.

Turning to the distributional facts, it can be seen that the phones [i, y, u] contrast in open syllables following /l/, as further illustrated in (6a). However, [ɨ] contrasts only with [u], since [ɨ] can only occur after dental and retroflex affricates and fricatives, whereas [i, y] cannot occur in these environments, as shown in (6b,c). Thus the phone [ɨ] does not contrast with either [i] or [y].

(6) a. [lu51] 'deer' [li51] 'power' [ly51] 'green'
b. [su51] 'pale' [sɨ51] 'four' *[si] *[sy]
c. [ʂu51] 'tree' [ʂɨ51] 'yes' *[ʂi] *[ʂy]

In sum, the phonetic and distributional facts show that /i, y, u/ contrast and are therefore members of three different phonemic categories. The high central vowel [ɨ], which does not have dental vs. retroflex variants in TM, contrasts with [u], but is in complementary distribution with both [i] and [y]. Thus it cannot be determined by the distributional facts alone which of these two phones it should be grouped with in a single phonemic category, if either. The psycholinguistic data to be discussed below will be crucial for supporting one hypothesis over another.

2.2.3. *Low Vowels and the Lower-Mid Vowels*

In TM, the low central vowel [a] occurs in open syllables, or in closed syllables before the coda consonants [n, j] as illustrated in (7a). The low back vowel [ɑ] only occurs before the velar consonants [ŋ, w], as illustrated in (7b). There is no phonetic [æ] in TM; the low front phone which occurs in the context of [j__n] is the lower-mid front vowel [ɛ]. However, there are some problems with analyzing the [ɛ] as part of the /a/ phoneme as opposed to including it in the mid-vowel phoneme.

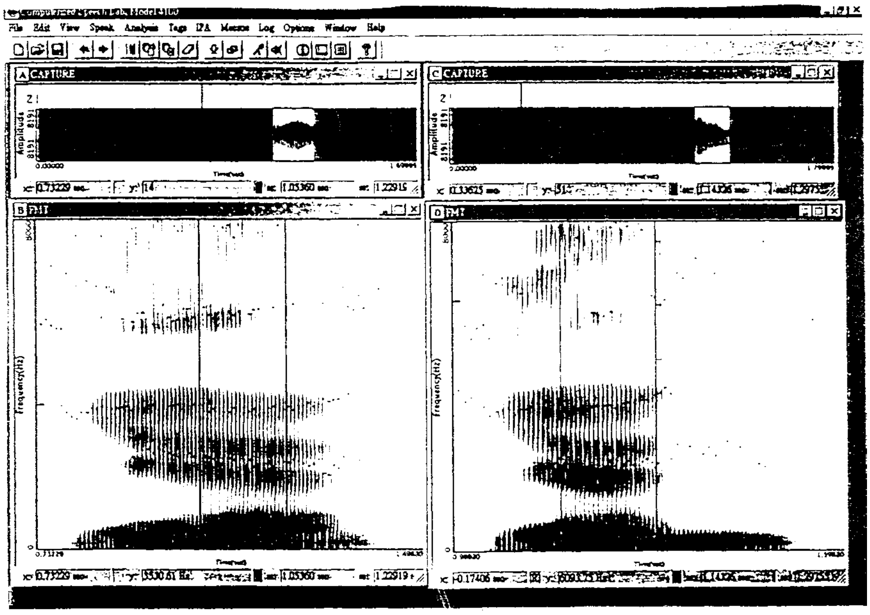
First, the phones [a] and [ɛ] contrast in open syllables following palatal glides, as shown in (7c); this is parallel to the fact that [a] and [ɔ] contrast in open syllables following the labial-velar glide, as shown in (7d).

(7) a. [ta55] 'to take (a bus)' [tan55] 'to serve' [taj55] 'stupid'
b. [tɑŋ55] 'to be' [tɑw55] 'knife'
c. [ja51] 'surprise' [jɛ51] 'leaves'
d. [wa21] 'roof' [wɔ21] 'I'
e. [jɛn55] 'smoke' [ɥɛn55] 'hatred' *[jan] *[ɥan]

Because [a] and [ɛ] contrast in otherwise identical environments, there is reason to believe that [ɛ] should not be grouped with the /a/ phoneme, at least according to standard theories of phonemic contrast.

As discussed above, Lin argues that there are two sources for [ɛ]. In open syllables, such as those shown in (7c), Lin analyzes the [ɛ] as a lowered version of [e] and thus an alternant of the mid-vowel phoneme. But the [ɛ] which occurs in the environment [j__n] or [ɥ__n], as illustrated in (7e), is argued to be derived from /a/, since the phone [a] cannot occur in this environment. Evidence for the latter came from the [r] suffixation process, which brings out a [a]–[ɛ] alternation in the same root, as discussed in Section 2.1 above. However, in TM this [r] suffixation process does not occur, and thus speakers would have no reason to believe that the [ɛ] in the environment [palatal glide + ɛ + n] is any different from the [ɛ] that occurs in open syllables following the palatal glides. Furthermore, there is no phonetic difference between the two manifestations of [ɛ], as shown in the example spectrograms in (8). These two utterances were spoken by the same male native speaker of TM who produced the spectrograms in (5) above.

(8) Spectrograms of [jɛ] vs. [jɛn]



An analysis of these two spectrograms shows that the steady-state vowel formants in [jɛ] are approximately F1: 605, F2: 2050, and F3: 3280; for the vowel in [jɛn] the formant values are F1: 610, F2: 2005, and F3: 3200.[3] Clearly these two vowels show no obvious phonetic differences which could be attributed to their being affiliated with different underlying vowel categories. Thus the second controversial point which will be looked at in detail in the psycholinguistic data below is whether or not the [ɛ] behaves as if it belongs to both the mid and low phonemes, or whether it shows a strong affiliation with only one of these two phonemes.

2.2.4. *Mid and Lower-Mid Vowels*

The occurrence of the mid vowel phones in TM depends on the status of adjacent segments, as shown in Table 2. The words in (9) below were chosen to illustrate some of the co-occurrence patterns of mid vowel phones with both onset and coda consonants.

(9)	a.	[ɤ51] 'hungry'		[kʰɤ51] 'lesson'
	b.	[ən55] 'mercy'		[pən51] 'stupid'
	c.	[oŋ55] 'old man'	[ow55] 'Europe'	[koŋ51] 'together'
				[kow21] 'dog'
	d.	[kwɔ35] 'nation'		
	e.	[kej21] 'to give'		[tjɛ55] 'father'
				[lɥɛ51] 'to plunder'
	f.	[kwej51] 'expensive'		[tɕjow51] 'uncle'
	g.	[joŋ51] 'to use'	[ɥoŋ51] 'to use'	[woŋ51] 'jar'
	h.	[jɛn35] 'along'	[ɥɛn35] 'round'	[wən51] 'to ask'

It can be seen from Table 2 that the mid vowel phones never occur after the palatal obstruents [ɕ, tɕ, tɕʰ]. The examples in (9a) illustrate the fact that the vowel phone [ɤ] occurs alone or in CV syllables where the C is neither a glide nor a labial consonant.[4] The examples in (9b) illustrate the fact that [ə] always occurs before the alveolar nasal [n], and can occur either with no onset or with obstruent onsets; glide onsets will be discussed below. While in some dialects of Mandarin [ə] can occur before the velar nasal [ŋ], in either onsetless syllables or syllables with onsets, in the dialect of TM spoken by the subjects in this study, all instances of [(CG)əŋ] have been neutralized to [(CG)ən]. In (9c), the examples show that [o] always occurs before [ŋ] or [w], in either onsetless syllables or in syllables with obstruent onsets; again, glide onsets will be discussed below. As seen in (9d), when [w] is in the prenuclear glide slot in codaless (C)GV syllables, the only mid vowel phone allowed in this position is [ɔ]. Example (9e) illustrates the fact that when there is a palatal glide [j] in the coda position, the mid vowel [e] occurs, but when there is a prenuclear palatal glide the vowel is realized as [ɛ]. (Recall that the rounded palatal glide [ɥ] occurs only in prenuclear, not postnuclear position.) When glides both precede and follow a mid vowel, the coda glide determines the quality; in (9f) the [e] in [kwej51] is conditioned by the coda [j], but the [o] in [tɕjow51] is conditioned by the coda [w]. When there is a glide onset and a nasal consonant coda, there are three different patterns: When the coda nasal is [ŋ], only the [o] mid vowel is allowed, as illustrated in (9g). When the coda nasal is [n], the front vowel [ɛ] occurs after the two front glides [j] and [ɥ], and the central vowel [ə] occurs after the labial-velar glide [w], as shown in (9h).

We summarize the allophonic variation of mid and lower-mid vowels as below. (Since no mid vowels occur following palatal obstruent onsets, this has not been indicated separately in the table.)

(10)

[ɤ] occurs in (C)V syllables where C is neither a labial obstruent nor a glide.
[ə] occurs in (C)V[n] or [w]V[n] syllables.
[e] occurs before [j] in CV[j] or [w]V[j] syllables.
[ɛ] occurs after [j, ɥ] in (C)GV([n]) syllables.
[o] occurs in (C)(G)VF syllables where the final is [ŋ] or [w].
[ɔ] occurs after [w] in (C)GV syllables.

In the psycholinguistic data we will look for evidence that all of these phones are associated with each other; in particular we will see whether [ɛ] is more closely related to the mid vowels vs. the low vowels.

2.2.5. *Summary*

In Table 3 we summarize all possible contrasts of surface vowels.

It can be seen from this table that there are many phones in Mandarin which contrast, and thus are candidates for being grouped into separate phonemic categories. On the other hand, there are many non-contrasting phones which are phonetically similar and could be candidates for being grouped in phonemes. For example, [ɨ] does not contrast with either [i] or [y]; [e, ɛ, ə, ɔ, o, ɤ] do not contrast with each other. The phone [u] does not contrast with [o, ɔ, ə, ɑ]; similarly [i] does not contrast with [e, ɛ]. Thus it is clear that the distributional facts alone do not resolve these phones into a unique phonological system. We now turn to a presentation of our psycholinguistic study, which will offer support to a specific resolution of this phonemic system.

3. Methodology and Data Collection

The use of speech errors, or 'slips of the tongue' to look at issues in linguistic theory as well as speech production planning models has a long

TABLE 3
All possible contrasts between surface vowels

	i	ɨ	y	u	e	ɛ	ə	ɔ	o	ɤ	a	ɑ
i		–	+	+	–	–	+	–	+	+	+	+
ɨ			–	+	–	–	–	–	–	+	+	–
y				+	–	–	+	–	–	+	+	–
u					–	–	–	–	–	+	+	–
e						–	–	–	–	–	+	–
ɛ							–	–	–	–	+	–
ə								–	–	–	+	–
ɔ									–	–	+	–
o										–	–	+
ɤ											+	–
a												–
ɑ												

'+' = contrast, '–' = no contrast.

history (Fromkin (1973, 1980), Dell (1980), Cutler (1982), Stemberger (1983), Garrett (1988), Levelt (1989), Baars (1992), to name a few). A slip of the tongue (hereafter 'SOT') is defined as a one-time error made during speech production planning, such that the speaker produces an utterance which is at odds with his or her intended utterance, differing from the intended utterance in terms of lexical choice, syntactic structure, or phonological structure. While most of this research has focused on errors from English, Dutch, or German, a few recent studies have looked into cross-linguistic data (Berg (1987), Wells-Jensen (1999)), and we have recently begun publishing on speech errors in Mandarin (Wan (1997, 1999, 2000, 2001, 2002), Wan and Jaeger (1998); see also Shen (1993), Chen (1993, 1999), Yang (1997)). SOTs have been shown to be invaluable evidence for the cognitive status of specific linguistic units and processes in specific languages, and this is our rationale for using the SOT methodology for looking at the vowel system of TM.

The current study is based on 238 SOTs involving vowels, selected from a corpus of approximately 3000 SOTs collected by the first author from native speakers of TM between 1995 and the present. Some of the errors were gathered from tape-recorded free conversations, during which the subjects did not know they were being recorded. After each conversation, the subjects were informed that the conversation had been recorded, and permission was obtained to use their data. Other errors were gathered following the 'pen and paper' method. That is, when slips occurred in situations where speakers were not being recorded (e.g., in normal conversations, during lectures, etc.), the errors were immediately written in a notebook. For each error the researcher recorded the complete utterance including self-corrections, and relevant contextual information; portions were written in IPA phonetic transcription when appropriate. Thus SOTs will be reported below in terms of the actual pronunciations subjects produced during the error utterance.

Errors were collected from approximately 80 different speakers. Subjects ranged from monolingual to trilingual, with TM as their first language and English and Taiwanese as their other language(s) if any. However, all the errors were collected when the speakers were conversing in TM; any errors which showed a bilingual influence were not included in the data set to be examined in this paper. Therefore we believe that these errors accurately reflect the processing involving in speaking TM.

The usual way to collect SOT data is to rely on the native-speaker linguist's intuitions as to what categories in the native language were heard by the native listener (Fromkin (1973)). This methodology is subject to some listener bias (Cutler (1982)). Furthermore, Mowrey and MacKay (1990) found that in SOTs induced in the laboratory by having speakers repeat 'tongue twisters' several times in succession, some phonetic differences between erroneously produced and intentionally produced consonants could be detected using electromyography. In the present study, one might argue that the perception by a native speaker that a syllable was spoken with a particular vowel is a more valid psycholinguistic measure than the actual phonetic properties of the utterance. However, the first author subjected several instances of erroneously produced vowels and the same vowels produced intentionally in the same environment to formant value analysis, and found no significant differences between the two (Wan (1999)). We thus assume that the data to be discussed below are sufficiently reliable to support our analyses.

In the earlier study by Wan (1999), it was found that SOTs in TM can occur at any stage of the speech production planning model (see Section 5.2 below for the details of our planning model). As indicated in the introduction, in this paper we are interested in errors which occur during the phonological planning stage, where underlying phonological representations are being assigned a surface phonetic form. There are two types of errors which can occur during this stage on which we are focusing. First, one vowel can be erroneously substituted for another. Second, consonants can be substituted, added or omitted next to a vowel, causing the vowel quality to shift; this can also involve a metathesis of a glide and vowel, causing the vowel quality to shift. These errors can occur contextually, that is, there can be a source for the error in the utterance itself; they can also be non-contextual, with no source in the error utterance. When an error is contextual, and the error occurs in a word spoken before the source of the error, this is called an 'anticipation'; if the error occurs in a word spoken after the source, it is a 'perseveration'. Consonants and vowels can also be exchanged.

In all of the examples which we will present below, the following notational format will be used: In the row headed by 'I', the intended TM utterance is given in IPA with the English glosses below; tones are given in tone numbers after the segments as follows: Tone 1 = 55, Tone 2 = 35, Tone 3 = 21,[5] Tone 4 = 51. Then in the row headed by 'E', the error utterance is given. Under the error utterance is a translation of the intended utterance. In every case, the phonological error resulted in an utterance which was meaningless, in the sense that it was either ungrammatical or uninterpretable, so we have not given translations of the error utterance. In the intended and error utterances, the 'source' unit(s) of the error (that is, the units which caused the error) are in boldface; the 'target' unit(s) (that is, the units planned for the intended utterance which were produced erroneously in the error) are underlined; and the actual 'error' (the elements spoken erroneously) are both boldfaced and underlined. In cases where a consonant has been erroneously added, substituted, or omitted next to a vowel, causing the vowel to be realized in a different surface form, both the vowel and consonant are underlined as the 'target', and both boldface/underlined as the 'error'. This is not meant to imply that the vowel was erroneously produced, but that the change in the vowel was made to accommodate the consonantal error (see Example (12) below). The two basic types of errors are as follows:

1. Phonological errors where one vowel is erroneously substituted for another, either due to the influence of another vowel in the same utterance, or non-contextually. If one vowel phone can be substituted for another, with the contiguous environment remaining the same, this proves that the vowel phones come from different underlying vowels. SOTs always produce sequences of segments which are legal in the language being spoken; if

the two phones were members of the same phoneme, substituting one for the other would produce an allophone disallowed in this environment, i.e., a disallowed phonotactic sequence, which never occurs. An example is given in (11).

(11) I: tʰɑŋ55-m**u**21 kɤ51-l**u**21-sɨ →
Tom Cruise
E: tʰɑŋ55-m**u**21 kɤ51-l**u**21-s**<u>u</u>**
'Tom Cruise'

In this error the vowel [u] from either the syllables [mu21] or [lu21] is perseverated and substituted for the vowel [ɨ] in the syllable [sɨ55], showing that [u] and [ɨ] are reflexes of differing vowel phonemes.

2. The second type of error which is of interest in this study is phonological errors in which the environment before or after the vowel is changed; that is, a contiguous segment is erroneously substituted, added, or omitted, or a vowel and contiguous segment metathesize. If changes occur in the vowel quality when these errors occur, then the two vowel phones can be taken to be conditioned variants of each other, as illustrated in (12).

(12) I: maj51-tɕi**ŋ**55-tʰ<u>a</u>21 →
Macintosh
E: maj51-tɕi**ŋ**55-tʰ **<u>ɑ</u>ŋ**21
'Macintosh'

In (12), the [ŋ] from the syllable [tɕiŋ55] is perseverated and added to the end of the syllable [tʰa21]; in this context the [a] is realized as [ɑ], suggesting that these two phones derive from the same phoneme.

Another possible way of interpreting such errors might be that when a consonant is erroneously substituted, added, or deleted, causing an illegal sequence of CV or VC, the vowel will be shifted to some other legal vowel phone in this environment, but that this is not necessarily an argument that the vowel produced in the error is an allophone of the intended vowel. And, as mentioned above, it certainly is a well-documented fact in the SOT literature that slips rarely produce an illegal sequence in the language (Fromkin (1973), Nooteboom (1973), Stemberger (1983)). The main problem with this argument, and particularly in terms of the example given in (12) above, is that in most cases there are a number of possible vowel phones which could occur legally in the environment which has been newly created by the error. So in the case of (12), the vowel phones [i, o, ɑ] could all legally occur in the vowel slot in the sequence [tʰ_ŋ]. If the shift in vowel were simply caused by the speaker 'cleaning up' the illegal string, we would expect that any of these vowel phones could be substituted for the intended phone [a] with equal frequency. But in fact in nearly every error where a vowel shifts quality when the consonant environment changes, the vowel alternates only with a single vowel phone (or in the case of the mid vowels, the same set of vowel phones) and not with other vowel phones which would be legal in this environment. In Section 4.3 below we document 85 cases in which the vowel phones [a] and [ɑ] alternated with each other in TM speech errors where the consonantal environment was shifted, and only one possible counterexample where [a] alternated with [ɛ]. This is extremely strong evidence that the analysis in terms of allophonic variation is the correct one, and certainly is in line with previous research on SOTs which documents that such allophonic variation is the norm (Stemberger (1983)).

An important methodological consideration has to do with making decisions on the classification of error types when more than one analysis is possible. Consider for example the error in (13).

(13) I: ta**n**55-ɕi**n**55 wɔ21 f<u>ɑŋ</u>35-tsɨ →
worry I house

E: ta**n**55-ɕi**n**55 wɔ21 f<u>**an**</u>35-tsɨ
'worry about my house'

One could make at least three arguments regarding this error. First, one could argue that the alveolar place feature from the two [n]s in the syllables [tan55] and [ɕin55] had perseverated and affected the velar place of [<u>ŋ</u>] in [f<u>ɑŋ</u>35], causing [<u>ŋ</u>] to be manifested as [n] and thus causing the vowel phone [ɑ] to front to [a]. Second, one could argue that the whole segment [n] had been substituted for [ŋ] also causing this vowel change. Finally one could argue that the rhyme [an] was perseverated from [tan55] and substituted for the rhyme [ɑŋ] in [fɑŋ35].

In making decisions on this type of error we use the 'Minimal Movement Principle' of Laubstein (1987), who argues that by far the most common unambiguous phonological error type is the substitution, addition, or omission of a single segment. Unambiguous feature errors are extremely rare in all languages, and Shattuck-Hufnagel and Klatt (1979) showed conclusively that most errors where one segment changes to a different segment are in fact errors of whole segments rather than features. Similarly, Fromkin (1973) argues that features are not 'independently controlled' in speech production, and therefore are not as susceptible as segments to error. Furthermore, all speech error studies have found that unambiguous errors involving syllables, rhymes, or other sub-syllabic units are extremely rare;

however, one segment is more likely to substitute for another if both target and source have identical adjacent segments (the 'repeated phoneme effect', Dell (1984)). Thus all of these principles would lead us to analyze errors such as that in (13) as a case where the segment [n] has substituted for the segment [ŋ], causing the vowel change; the fact that both nasal codas were adjacent to the low vowel phoneme simply makes this nasal segmental substitution error more likely to occur because the codas share a common adjacent environment.

4. Findings

4.1. *Overall Findings*

In looking at our data, we will ask the following questions.

1 Which vowel phones can substitute for each other in SOTs, without necessitating any other changes in the syntagmatic string?
2. When the environment contiguous to a vowel phone changes, which vowel alternations occur?
3. If some vowel phones are judged to be contextual variants of the same phoneme, is one of the allophones the more basic variant? In making this decision, we are making a decision about the structure of the underlying vowel system of TM.
4. Finally, do our results concur with any of the four analyses of Mandarin vowels presented in Section 2.1 above, or do they require us to develop a new analysis appropriate for the vowel system of TM?

We will present general results here, and then discuss the various sets of vowel phones by vowel height in the following sections.

Our data revealed 41 errors in which one vowel substituted for another, with no other change in the error word, as illustrated in (11) above. Furthermore there were another 15 errors in which a glide, which we assume is derived from a high vowel, substituted for a vowel but then was realized as its vocalic alternant, or vice versa. An example is given in (14).

(14) I: **ɥ**ɛn21-l<u>i</u>35 →
far-away
E: **ɥ**ɛn21-l**<u>y</u>**35
'far away'

In this error, the front rounded vowel [y] is realized as its glide alternant [ɥ] in the first syllable [ɥɛn21] since it is adjacent to a non-high vowel. However, when it substitutes for the vowel [i] in the second syllable [li35], it then takes on its vocalic alternant, so that the error is produced as [ly35]. If we combine all the vowel-vowel substitutions with the vowel-glide substitutions, we find the pattern of vowel substitution shown in Table 4 (N = 56). This table will be discussed in detail below.

Second, we tabulated every error in which one vowel phone shifted to another due to a change in the environment in which the vowel occurred. The results are shown in Table 5 (N = 155). Again we will discuss this table in detail in the following sections.

Finally, there were 27 other errors in which vowels and glides interacted. In some cases a vowel was added or omitted, causing a vowel to become its glide alternant or vice versa. In other cases a vowel or glide was added in a new location, but was required to take on the opposite syllabicity due to its new syntagmatic status. These errors included 15 involving the [i-j] alternation, 8 involving the [w-u] alternation, and 1 involving the [y-ɥ] alternation, which is given in (15). There were also 3 errors where a vowel substituted for a coda nasal, causing the glide alternant to occur (two involving [w] and one involving [j]; [ɥ] does not occur in coda position so is not eligible for this type of error).

TABLE 4
The number of errors in which one vowel phone substituted for another

	i	ɨ	y	u	e	ɛ	ə	ɔ	o	ɤ	a	ɑ
i			19	5			1			1	1	
ɨ				7								
y				7								
u										2	1	
e											2	
ɛ												
ə											4	
ɔ												
o												6
ɤ												
a												
ɑ												

TABLE 5
The number of errors in which a vowel phone shifts to a different vowel quality when the adjacent environment changes

	i	ɨ	y	u	e	ɛ	ə	ɔ	o	ɤ	a	ɑ
i		9										
ɨ												
y												
u												
e						3	6	7	1	3		
ɛ							11	1	4	1		1
ə								3	3	2		
ɔ									5	8		
o										2		
ɤ												
a												85
ɑ												

(15) I: t^hjɛn55 ɕɥɛn35 ti51 tʂwan21 →
sky turn ground turn
E: t^hjɛn55 ɕyn35 ti51 tʂwan21
'The whole world is turning over.'

In (15), the vowel [ɛ] was deleted in dissimilation from the vowel [ɛ] in the preceding word; when the /y/ found itself in syllable nucleus position, the vowel [y] rather than the intended glide [ɥ] was produced. We take these errors to be extremely strong evidence that glides derive from high vowel phonemes. One other possible interpretation would be that when the vowel was erroneously deleted, the illegal sequence *CGV was produced, with no vowel nucleus; in this case the glide was replaced by some vowel by a sort of ad hoc on-line processing constraint. However, if this sort of process was common, we would expect to find that there would be no consistent relationship between the glide and the vowel which substituted for it (or vice versa), when in fact in every error of this sort the relationship between [i]-[j], [y]-[ɥ], and [u]-[w] was consistent (N = 42). Furthermore, there were no errors in our corpus which involved the creation of an illegal phonotactic sequence; for example, there were no errors where a vowel was deleted, producing a syllable such as *CF with no nucleus. This suggests that the processing mechanism is treating the glide as if it is a nucleus in underlying structure, so if there are two potential nuclei in sequence in a syllable, the non-high vowel may be deleted. The gliding which appears on the surface is a later process which occurs after the SOTs have occurred, if a sequence of high + non-high vowel (or vice versa) remains in the surface form. For all these reasons we will take as a given in the remainder of this paper the hypothesis that glides are derived from high vowel phonemes, which is consistent with the theories of Lin (1989) and Wu (1994) discussed above. We will now turn to a detailed discussion of groups of vowel phones, starting with the high vowels.

4.2. *High Vowels*

Following the above guidelines, the high vowel errors will be interpreted in terms of the following questions:

1. In substitution errors involving high vowels [i, y, ɨ, u], are these phones all equally likely to be substituted for each other?
2. When the environment before or after these four high vowel phones changes, do any of these phones alternate with each other?
3. Are there any errors in which the vowel [y] shows evidence of being underlyingly a sequence of [u] and [i], as in Wu's (1994) proposal? That is, are there errors in which it appears that only part of this sequence participates?
4. What are the underlying high vowel phonemes of TM?

Looking first at Question 1, we found 25 cases involving substitution among the high vowels [i, u, y, ɨ], and another 13 substitution errors involving interactions of glides with high vowels. Of the 25 simple vowel substitutions, 10 involved the interaction between [i] and [y], as exemplified in (16); 9 of the vowel/glide interaction errors also involved [i/j] and [y/ɥ].

(16) I: p^{h}**i**ŋ35-tɕyn55 p^{h}**i**ŋ35-tɕy̠n55 →
average average

E: p^{h}**i**ŋ35-tɕyn55 p^{h}**i**ŋ35-tɕ**i̠**n55
'pretty much average'

In Example (16), the vowel [i] is perseverated and substituted for the vowel [y], but since [tɕin] is an allowable sequence in TM, no further changes in the word are required. The next most common substitution (N = 7) involved an interaction between [u] and [ɨ], as illustrated in (17).

(17) I: tʰa55 ʂɨ51 tswɔ51 tsɨ̲51-tʂ**u**51-tsʰan55 →
he is do cafe

E: tʰa55 ʂɨ51 tswɔ51 ts**u̲**51-tʂ**u**51-tsʰan55
'he has a café'

In Example (17), the vowel [u] is anticipated and substituted for the vowel [ɨ], producing the legal sequence [tsu]. An equally common substitution involved [u] and [y]; 5 vowel substitution errors and 2 vowel/glide errors involved this pair of phones, as illustrated in (18).

(18) I: lu̲51 tɕ**y**51 wɔ21-mən xɑw21 tɕin51 →
deer distance us very close

E: l**y̲**51 tɕ**y**51 wɔ21-mən xɑw21 tɕin51
'The deer is very close to us.'

In this error, the vowel [y] is anticipated and substituted for the vowel [u]. The remaining 3 vowel substitution errors and 2 vowel/glide errors involved the pair [i] and [u], as exemplified in (19).

(19) I: pu̲35 jɑw51 tʰjɑw35 pʰ**i**35 →
not want naughty

E: p**i̲**35 jɑw51 tʰjɑw35 pʰ**i**35
'don't be naughty'

In Example (19), the vowel [i] is anticipated and substituted for the vowel [u]. We found no cases where [i-ɨ] or [y-ɨ] substituted for each other. This is to be expected, since [ɨ] only occurs after dental and retroflex obstruents, and [i, y] never occur in these environments.

These data suggest that /i, y, u/ are separate phonemes, and that [ɨ] and [u] belong to distinct phonemes. Table 4 also shows that there are vowel substitution errors involving [i]-[ə, ɤ, a], and [u]-[ɤ, a], verifying that these phones belong to separate phonemes. However, the vowel substitution data do not tell us whether [ɨ] should be grouped with either [i], or [y], or neither.

In order to look into this issue, we need to look at errors where the environment changes, causing a concomitant change in the vowel. However, it turns out that unambiguous errors of this sort are difficult to find for high vowels, since the consonantal environment which is relevant for conditioning the occurrence of these three high front or central vowel phones, palatalization, could be considered to be a function of the vowel quality. Thus we are in danger of falling into a circular reasoning pattern. Consider the following examples:

(20)	a. I:	tjɛn51-t**ɕ**i55	pwɔ35-<u>s</u>ɨ51	pan55 →
		electrical-engineering	Ph.D.	class
	E:	tjɛn51-t**ɕ**i55	pwɔ35-**ɕ**i51	pan55
		‘Ph.D. program in the Department of Electrical Engineering’		
	b. I:	koŋ55-**ts**wɔ51	<u>tɕ</u>i55-xwej51 →	
		job	opportunity	
	E:	koŋ55-**ts**wɔ51	**tsɨ**55-xwej51	
		‘job opportunity’		

Wan (1999) argued that these errors show either a substitution of the feature of palatalization (as in (20a)) or a deletion of the feature of palatalization (as in (20b)) in the consonant, due to the influence of other palatal or non-palatal consonants in the utterance. In the previous analysis it was not claimed that (20a) showed a substitution of [ɕ] for [s], since these two segments are considered to be derived from the same underlying phoneme; similarly in (20b), [tɕ] and [ts] are considered allophones of the same phoneme. This grouping of the dentals and palatals into single phonemes was argued for due to the fact that whenever the palatal feature is deleted (as in (20b)), the consonant reverts to the dental place of articulation; similarly, if a palatal glide is inserted after a dental obstruent, it is realized as a palatal obstruent. However, one could argue that these are vowel substitution errors. In (20a), one could claim that the [i] from [tɕi55] was perseverated and substituted for the [ɨ] in [sɨ51]. But in (20b) it would be difficult to argue that there is a substitution of the vowel [ɨ] for [i], since there is no [ɨ] elsewhere in the utterance which could serve as the source of the error, and such substitutions are overwhelmingly contextually motivated. Therefore, if we accept the Wan (1999) analysis of these errors for the sake of argument, then our corpus yields 9 errors in which the consonantal environment shifts from dental or retroflex to palatal (as in (20a)) or vice versa (as in (20b)), and in every case, the high front unrounded vowel [i] alternates with the high central unrounded vowel [ɨ]. There are no cases in our errors where the environment shifts and [ɨ] shows an alternation with the other palatal vowel, [y]. Specifically, there are no cases where the palatal environment is added, such as in (20a) above, and the intended vowel [ɨ] is realized as [y], even though the sequences [ɕy], [tɕy] and [tɕ^{h}y] are legal. Furthermore, there are no errors in our data where the palatal environment is removed before [y]; we suspect that this is because this would result in an illegal sequence of a non-palatal consonant + /y/, and /y/ does not have a non-palatal allophone to revert to. Two other examples involving consonant addition may serve to strengthen this argument.

(21) a. I: ʐ**ən**35 jɑw51 <u>tsɨ</u>55-tsu35 →
people need content
E: ʐ**ən**35 jɑw51 **<u>tɕin</u>**55-tsu35
'People need to be content with what they have'

b. I: <u>i</u>35-tiŋ51 jɑw51 kʰan51 **s**an55-ʂɨ35 fən55 ʂu55 →
must need study 30 minute book
E: **<u>s</u>**ɨ35-tiŋ51 jɑw51 kʰan51 **s**an55-ʂɨ35 fən55 ʂu55
'(you) must study for 30 minutes'

In (21a), the nasal [n] is perseverated from [ʐən] and added after the vowel [ɨ] in [tsɨ55-tsu35].[6] As shown in Table 2, the vowel [ɨ] can only occur in open syllables in TM, and thus sequence *[tsɨn] is not legal; therefore the vowel shifts to its legal allophone [i], which can occur before [n]. This causes the preceding affricate to palatalize from [ts] to [tɕ]; (note that the sequence [tɕyn] is also a legal sequence, but does not occur in this error). In (21b) we find the opposite case: the [s] from [san55] is anticipated and added into the onset slot of [i35]; since the sequence of dental fricative or affricate + [i] is not allowed in TM, the vowel [i] shifts to the appropriate allophone [ɨ] after [s]. Thus there is clear evidence that the high central vowel [ɨ] should be grouped together with the high front unrounded vowel [i] into a single phoneme, since it is in complementary distribution with this phone, and these two phones alternate when the consonantal environment shifts. There is no reason to consider grouping it with the rounded vowel [y]; although [y] and [ɨ] do not contrast, they also never alternate in any of these speech errors.[7]

Moving on to Question 3, we find that there are no cases where the vowel [y] splits, and only one part is involved in an error. In every case where [y] is involved in an error, it behaves identically to every other vowel phone in the data, i.e., like a whole unit rather than a sequence of units. This is illustrated in Examples (15), (16), and (18) above. More evidence for the fact that [y] is not derived from a sequence of /ui/ is that these three vowels [i, y, u] freely substitute for each other as whole segments; we have no cases where, for example, a sequence of vowels substitutes for a single vowel, or vice versa. We do find it a bit odd that we found only one error where [y] and [ɥ] alternate with each other when the environment shifts, as we found several for [i-j] and [u-w]; but this may be due to the fact that the front rounded vowel/glide phoneme is far less frequent than the other two high vowel phonemes. Finally, we are in principle opposed to an analysis which derives the high front rounded vowel/glide from a sequence of underlying high vowels, since in all other theories the only possible sequences of vowels in phonological representations must involve a high with a non-high vowel. This proposal makes a very different claim about the possible underlying structures of morphemes in Mandarin, one which would require more supporting evidence. We therefore see no reason to consider the vowel /y/ as anything other than a phoneme in its own right.

In summary, our findings show that the vowels [i, y, u] and the vowels [ɨ, u] belong to separate phonemes, since they can be substituted for each other. The vowels [i, ɨ] alternate with each other when the preceding environment alternates between palatal and non-palatal, showing that they are allophones of a single phoneme. There is no evidence to support the claim that the vowel [y] is derived from a sequence of /ui/, since [y] never splits into components in errors, and can substitute for both [u] and [i] in errors.

The final remaining question is: What is the basic phonemic system of the vowels of TM? Because the phone [i] occurs in far more environments than [ɨ], we will assume that [i] is the basic variant, which gives us the universally more common vowel system /i, y, u/ (as compared to */y, ɨ, u/).

Comparing our findings to the four theories discussed in Section 2.1 above, we find that our analysis is different from all four in that we claim that there is a single high central unrounded vowel [ɨ] in TM rather than two distinct 'apical' vowels. Other than this difference, however, our analysis is most similar to the 5-vowel system of R. Cheng (1966), who considers /i/, /y/, and /u/ to be distinct phonemes, and derives the apical vowels from /i/. We differ from C. Cheng who considers /i/ and /ɨ/ to be distinct phonemes. We also differ from Lin who derives the apical vowels from an epenthesis and assimilation process. Finally, we differ from Wu in that, even though she considers /i, u/ to be phonemes, and derives the apical vowels from /i/, she considers [y] to be derived from a sequence of /ui/; we find no psycholinguistic evidence for this latter analysis.

4.3. *Low Vowels*

Errors involving the low vowels will be interpreted in terms of the following questions:

1. Do the vowel phones [a] and [ɑ] ever substitute for each other in an error, with no other changes?
2. If the environment contiguous to one of these vowel phones changes, does the central phone [a] shift to the back phone [ɑ] and vice versa?
3. Are there any cases where either of the two low vowel phones [a, ɑ] alternate with the lower-mid front vowel [ɛ] when the environment changes, as would be predicted by the four analyses discussed in Section 2.1 above?
4. If [a] and [ɑ] alternate, which is the basic variant?

Looking first at Question 1, we find no cases in our data where [a] substituted for [ɑ] or vice versa with no other changes in the word, as shown in Table 4. It can be seen from this table that there are substitution errors involving [a]-[i, u, e, ə], and [ɑ]-[o], verifying that these high and mid vowel phones contrast with the low vowel phones, and thus should not be grouped into phonemes with [a] or [ɑ].

Table 5 shows a clear answer to our second question: We have 85 errors in which the environment contiguous to one of these two low vowels changes, causing the vowel to shift to the other variant. Seventy-three of these errors involve the substitution of one segment for another in the coda slot following one of these phones; in 30 cases, the non-velar codas [n, j] are substituted for the velar codas [ŋ, w], causing [ɑ] → [a], as illustrated in (22), and in 39 cases the reverse occurs, as illustrated in (23); in the remaining 4 cases the nasals [n] and [ŋ] exchange with each other, causing the vowels to shift.

(22) I: jɛ**n**21-kw<u>ɑŋ</u>55 →
eye-light

E: jɛ**n**21-kw<u>**an**</u>55
'judgment'

In Example (22), when the nasal [n] is perseverated and substituted for the nasal [ŋ], the vowel [ɑ] changes to [a].

(23) I: n<u>an</u>35-t^hi**ŋ**55 sɨ21 lɤ →
hard-listen dead

E: n<u>**ɑŋ**</u>35-t^hi**ŋ**55 sɨ21 lɤ
'very bad music'

In Example (23), when the nasal [ŋ] is anticipated and substituted for the nasal [n], the vowel [a] changes to [ɑ].

Addition and omission errors give further evidence regarding Question 2. Our data contain 9 errors in which the vowel [a], which is in an open syllable in the target word, is realized as the back [ɑ] when the velar codas [w, ŋ] are erroneously added after the vowel, as exemplified in (24).

(24) I: l**w**ɔ35-m<u>a</u>21 →
Rome
E: l**w**ɔ35-m<u>**ɑw**</u>21
'Rome'

In Example (24), the glide [w] is perseverated and added after [a], and [a] changes to [ɑ].

Similarly, we found 3 errors in which [ɑ] is realized as [a] when the coda glide [w] or nasal [ŋ] is erroneously omitted, as illustrated in (25).

(25)	I: tsu51-tɕj<u>ɑw</u>51	pi21	sej35	to**w**55	ta51 →	
	assistant	than	anyone	much	powerful	
	E: tsu51-tɕj<u>**a**</u>51	pi21	sej35	to**w**55	ta51	
	'The assistant is more powerful than anyone else'					

In Example (25), the glide [w] in [tɕjɑw51] is omitted in dissimilation from the [w] in [tow], causing the [ɑ] to occur in an open syllable so that it must be realized as the central [a] variant. From these substitution, addition, and omission errors we find very strong evidence that [a] and [ɑ] are contextual variants of the same phonemic category, as would be predicted by all four theories discussed above.

However, Question 3 is more problematic. First, we found no errors in our corpus where [ɛ] substituted for [a] or [ɑ]. One could argue that this supports the claim that [ɛ] could be grouped with the low vowels into one phonemic category. However, when we look at errors where the environment changes and vowels shift quality, we find a very different pattern. In every case except one, when the environment contiguous to [ɛ] changes so that this is no longer a legal phone in the new environment, the [ɛ] shifts to one of the other mid vowels; similarly, if the environment contiguous to one of the mid vowels is changed so that [ɛ] is the appropriate allophone, all the other mid vowels shift to [ɛ] (N = 20); these figures are shown in Table 5, and will be discussed in detail in Section 4.4 below. This includes 15 errors in which the erroneous environment is exactly the environment where Lin (1989) argues the [a-ɛ] alternation should occur. Consider the following example.

(26) I: tsan51-tsa**j**51 i55-p<u>jɛ</u>n55 →
stand in one-side

E: tsan51-tsa**j**51 i55-p<u>ə</u>n55
'(People are) standing together to one side'

In this example, the prenuclear glide [j] is omitted in the syllable [pjɛn55], in dissimilation from the [j] in [tsaj51]. Because there is no longer a palatal prenuclear glide in this syllable, the vowel must shift to another phone which is allowed in the sequence [p__n]. The crucial factor is that both [pən] and [pan] sequences are allowed in TM, and yet the error the speaker actually produced was [pən], which argues strongly that the underlying phonological representation of the target word is /piən/ and not /pian/. It is clear from the unambiguous data discussed above regarding the [a-ɑ] alternation, that the default production of an error word will maintain the same phoneme as in the target word if possible, when the environment next to the phoneme changes, even if some other phoneme can occur there. Otherwise we might expect, for example, that the [a] in [nan35] in Example (23) above might change to an [o] when the [ŋ] substitutes for the coda [n], since the sequence [noŋ35] is legal in TM. However, this sort of random vowel change after a consonant substitution doesn't occur in our data. Thus, since we have 20 errors in which [ɛ] alternates with other mid-vowel phones when the environment is erroneously changed, and only one error in which this does not occur, this is very strong evidence for the affiliation of [ɛ] with the mid vowel phoneme in TM and not the low vowel phoneme.

The one counterexample is given in (27).

(27) I: tʂo**ŋ**55-tɕj<u>ɛn</u>55 tɕjow51-ɕi**ŋ**35 lɤ →
middle just okay

E: tʂo**ŋ**55-tɕj<u>**ɑŋ**</u>55 tɕjow51-ɕi**ŋ**35 lɤ
'(putting it) in the middle is okay'

In this error the consonant [ŋ] was either anticipated or perseverated and substituted for the consonant [n] in the syllable [tɕjɛn], giving the illegal sequence *[tɕjɛŋ], since [ɛ] doesn't occur before the velar nasal. In this case, the vowel [ɛ] was realized as the low back [ɑ] in the error. If we wanted to argue that the [ɛ] derives from the mid vowel phoneme, we would have expected that the error phone would be [o], since a sequence of [tɕjoŋ] is legal for at least some TM speakers, as in the word [tɕjoŋ21], an archaic word meaning 'embarassed'; however, this is the only word with this sequence in TM. If the speaker who made the error in (27) knew this word, then we would have to conclude that there may be some affiliation of [ɛ] with /a/ for this speaker, since in this one case the two alternated and the expected [o] did not occur. However, if this speaker did not know this word, then this would be a simple case of the speaker substituting the [n] with [ŋ], and then producing the closest legal sequence possible, which would be [tɕjɑŋ]; in this case this error would not be as strong a counterexample to the claim that [ɛ] and [ɑ] are not affiliated with each other. Even if this were a counterexample, we would argue that because there were 20 errors which showed an alternation between [ɛ] and the other mid vowels, and only this one case showing the [ɛ]-[ɑ] alternation, we would argue that there is much more evidence for the inclusion of [ɛ] in the mid vowel than the low vowel phoneme in general. We will return to this point below in our discussion of mid vowels.

Finally, turning to Question 4, there is clear evidence that [a] is the basic variant of the phoneme which includes the [a-ɑ] alternants. This is because whenever the coda consonant is omitted following one of these two phones, only the phone [a] is realized; since this is the phone which occurs when there is no conditioning environment (i.e., no environment which would cause a backing or fronting assimilation), it must be the basic variant. This phenomenon was illustrated in Example (25) above.

In summary, there are 85 errors in our corpus in which the [a] and [ɑ] phones alternate; these alternations were caused by the substitution, omission, or addition of consonants after the vowel, such that only [a] occurred before [n, j] or syllable finally, and only [ɑ] occurred before the velar consonants [w, ŋ], clearly a place of articulation assimilation. Because only [a] occurs in open syllables with no conditioning environment, the low vowel phoneme is considered to be /a/. Finally, we found no evidence that [ɛ] should be considered a member of the low vowel phonemic category; we will return to this point in the next section.

4.4. *Mid Vowels*

Data regarding the grouping of mid vowels will be interpreted in terms of the following questions.

1. Can any of the mid vowel phones [e, ɛ, ə, ɔ, o, ɤ] be substituted for each other without any other change in the contiguous environments?
2. Which of these phones alternate with each other when the contiguous environment is changed?
3. What are the details of the evidence discussed in Section 4.3 that [ɛ] should be grouped with the mid vowels rather than the low vowels?
4. How many vowel phonemes are represented by these mid-vowel phones, and what is/are the basic variant(s)?

Regarding Question 1, we found no cases in our data where any of these mid vowel phones substituted for each other with no other changes in the syllable. Table 4 shows the following substitutions involving mid vowels: [ə]-[i, a], [e]-[a], [o]-[ɑ], [ɤ]-[u, i], indicating that mid vowels belong to distinct phonemes from both the high vowels [i] and [u], as well as the low vowels [a] and [ɑ].

In our data we have 60 errors which involve alternations among [e, ɛ, ə, ɔ, o, ɤ], as shown in Table 5, which provide evidence for Question 2. In 19 cases, one consonant substitutes for another either in the onset or coda of the target word, causing the target mid vowel to shift to one of the other phones. This includes 6 errors involving the [ə]-[e] alternation as exemplified in (28), 5 errors involving [ə]-[ɛ] as exemplified in (29), 4 errors involving [ɛ]-[o] ((30)), 3 errors involving [ə]-[o] ((31)), and one error involving [ɛ]-[ɔ], given in (32).

(28) I: **n**e̲j̲51-ljɛ**n**51 →
reserved

E: **nə<u>n</u>**51-ljɛ**n**51
'reserved'

In Example (28), the dental nasal [n] was either perseverated or anticipated and substituted for the glide [j] in the coda of the target word [nej51]. Since the [e] allophone can only occur before the glide [j], the vowel is realized as [ə] which is the correct allophone before [n].

(29)	I:	p**u**35	<u>ɥɛ</u>n51	t**u**35-s**u**55 →
		not	want	study
	E:	p**u**35	**<u>wə</u>**n51	t**u**35-s**u**55
		'(she) doesn't want to study'		

In this error, the vowel [u] was perseverated or anticipated and substituted for the [ɥ] in the onset slot of the target word [ɥɛn51]; in this context the /u/ is realized as the glide [w]. Since [ɛ] cannot occur after [w], the correct allophone for this environment, [ə], was spoken in the error. Notice that this is another case where [a] would have been legal in the error utterance, since the utterance [wan] is legal in TM; however, in this error what was actually spoken was the vowel [ə].

(30) I: mi**n**35-ɕɥ<u>oŋ</u>35 →
Ming-Xiong

E: mi<u>n</u>35-ɕɥ**<u>ɛn</u>**35
'Ming-Xiong (county name)'

In Example (30), the dental nasal [n] is perseverated and substituted for the velar nasal in the coda of the target word [ɕɥoŋ35]. Since [o] can occur only before the velar nasal, not the dental nasal, the vowel shifts to [ɛ], which is the correct allophone in the context [ɥ__n].

(31) I: tʰa55 na**n**35-sow51 sɨ21-lɤ →
he feel bad dead

E: tʰa55 na**n**35-s**ən**51 sɨ21-lɤ
'He felt extremely bad'

In Example (31), the dental nasal [n] is perseverated and substituted for the [w] in the coda slot of the target word [sow51]. Since, as discussed for Example (30), the sequence *[on] is disallowed, the vowel shifts to the [ə], which is the correct allophone in the context [s__n].

(32) I: **w**ɔ21 ɕɥɛ35-tɕjɛ21 →
my schoolmate

E: **w**ɔ21 **swɔ**35-tɕjɛ21
'My schoolmate'

The simplest explanation for the error in (32) is that the [w] from [wɔ] was perseverated and substituted for the [ɥ] in [ɕɥɛ]; this caused the palatal fricative [ɕ] to revert to its dental allophone [s]. In the context of [sw__$], where '$' marks the syllable boundary, the only mid-vowel alternate which can occur is the [ɔ], which is what was spoken in this error. (One could also argue that the prenuclear glide and vowel from [wɔ21] substituted as a unit for the [ɥɛ] sequence in the target word, but this sort of consonant cluster split in an error is extremely rare.)

Several other types of errors add more support to the claim that this set of 6 phones should be grouped together into a single phoneme. First, we have 25 errors in which a consonant is added either before or after the vowel, and the vowel shifts to the appropriate allophone. This includes 7 cases of a [ɔ]-[ɤ] alternation, 6 cases of a [e]-[ɔ] alternation, and 2 cases each of the following alternations: [ə]-[ɛ], [ə]-[ɔ], [ə]-[ɤ], [e]-[ɤ], [o]-[ɔ] and [o]-[ɤ]. Two examples will illustrate this process.

(33) I: tʰa55 tɕ**j**a55 pən21-la**j**35 tʂu51-tsaj51 tɕja55-li21 →
he house originally live-in Jiali

E: tʰa55 tɕ**j**a55 p**jɛ**n21-la**j**35 tʂu51-tsaj51 tɕja55-li21
'The house he originally lived in was in Jiali'

In this error, the glide [j] is anticipated/perseverated and added into the prenuclear glide slot of the target word [pən21]. Since the schwa cannot occur after [j], the vowel shifts to the correct allophone for this sequence [pj__n], which is [ɛ].

(34) I: t**w**ɔ55-pan51 tsɤ51-sɨ35 →
most this time

E: t**w**ɔ55-pan51 ts**ow**51-sɨ35
'usually at this moment'

In this error the glide [w] is perseverated from the onset of the first word and added into the coda slot after the vowel in the target word [tsɤ51]; since [ɤ] can only occur in open syllables, the vowel is realized as the correct variant for this context [ts__w], which is [o].

Besides the substitution and addition errors, there are 7 omission errors in our corpus in which a consonant before or after the vowel is omitted, causing the vowel to shift to a different alternant. These errors include 4 instances of the [ɛ]-[ə] alternation, as illustrated in (26) above, where the omission of the prenuclear [j] causes [ɛ] to revert to [ə]. Other alternations produced by omission include one each of [ɛ]-[ɤ], [e]-[ɤ], and [ɔ]-[ɤ]; an example is given in (35).

(35) I: xej55-tɕ^{h}**j**ɑw35 p^{h}a**j**35 →
Heiqiao brand

E: x**ɤ**55-tɕ^{h}**j**ɑw35 p^{h}a**j**35
'Heiqiao brand (a brand of sausages)'

In this error, the glide [j] was deleted from the coda location in the target word [xej55] in dissimilation from the other palatal glides in the utterance. This produced the sequence *[xe], which is illegal since [e] only occurs before the palatal glide. Thus the vowel shifted to the correct variant [ɤ], which occurs in the context [x__$].

In addition to the fairly straightforward errors discussed above, there were 5 errors which involved the metathesis of a vowel and glide, causing a different vowel to surface. This included 3 cases of the [ɛ]-[e] alternation, as illustrated in (36), and two cases of the [o]-[ɔ] alternation.

(36) I: mej35 li21-mɑw51
not polite

E: m**jɛ**35 li21 mɑw51
'not polite'

In this error the sequence of underlying vowels in the target word is /(mid vowel) i/. In the target utterance, the /i/ is realized as the postnuclear glide

[j], and the mid vowel is realized as the correct variant before the palatal glide, [e]. In the error word, the two vowels have metathesized, causing the sequence /i (mid vowel)/; in this case the /i/ is realized as the prenuclear glide [j], and the vowel takes on the correct allophone for the sequence [j__$], which is [ɛ]. Note that this is a case where Lin would agree that the [ɛ] is derived from the mid vowel, since it is in an open syllable. This example is exactly analogous to the 2 cases where /(mid vowel) u/ [ow] metathesizes to produce [wɔ].

Finally, there were 4 other errors which involved combinations of addition, substitution, omission, and metathesis of consonants, and which caused mid vowels to shift to different realizations. This included one each of the following four alternations: [ɔ]-[ə], [ɔ]-[e], [o]-[e], and [o]-[ɔ]. Two examples are given below.

(37) a. I: **s**wɔ35-i21 →
so
E: **sej35**
'so'

b. I: tsaj51 t^hu35-su55-k**w**an21 tən21 →
in library wait
E: tsaj51 t^hu35-su55-k**w**an21 t**wɔ**21
'waiting in the library'

The error in Example (37a) is called a 'telescoping' error, since elements from sequential syllables are omitted and the syllables are collapsed into one. In this case the [w] in the target word [swɔ] is omitted, and the [i] from the second syllable is resyllabified with the vowel [ɔ] of the first syllable. The vowel [i] now becomes a postnuclear glide [j], but the sequence *[sɔj] is illegal, so the [ɔ] shifts to the correct allophone for this sequence, which is [e]. In (37b), it appears that the [w] from the source word [kwan] has been perseverated and added into the prenuclear glide slot of the target word [tən]; furthermore the [n] in the target word has been omitted in dissimilation from the [n] in the source word. This leaves the sequence *[twə], and thus the [ə] is realized as the correct allophone for the sequence [w__$], which is [ɔ].

A glance at Table 5 shows that every one of the six mid-vowel phones alternates with every other one of these phones in at least one error, and that overall there are 60 errors showing these alternations. This is very strong evidence that these 6 phones should be grouped into a single phoneme. Moving on to Question 3 above, the status of [ɛ], these data make the answer very clear. In these mid vowel errors there are 20 cases where the vowel [ɛ] alternates with other mid vowel phones, and [ɛ] enters into alternations with all 5 of the other phones. In 15 of these cases, the [ɛ] occurred in the environment [j__n], which is the environment in which Lin argues the [ɛ] is derived from /a/. In this way [ɛ] behaves like any other mid vowel phone, and shows no behavior which would set it apart from the other phones. Although there was one example discussed in Section 4.3 above where [ɛ] alternated with [ɑ], this is not convincing evidence that [ɛ] should be considered a member of the /a/ phoneme, since there is much more evidence for its association with the other mid vowels. We conclude that in the dialect of Mandarin under study here, Taiwan Mandarin, the phone [ɛ] is a member of the mid vowel phoneme.

The grouping of these six allophones into a single phoneme allows us to account for some data which is otherwise difficult to explain. In our data we have 14 errors which we classified as 'non-contextual', that is, where there was no obvious source for the error in the utterance. Non-contextual errors are typically quite rare, since most phonological errors are caused by the substitution or addition of some other segment in the planned string, the exchange or metathesis of two planned segments, or omission in dissimilation from another segment in the string. However, if we adopt the hypothesis that these six phones derive from the same underlying phonological category, then 10 of these non-contextual errors involving mid vowels become contextual, since the source could be said to be a different allophone of this phoneme which does occur in the utterance. Two examples are given in (38).

(38) a. I: k^haj55 w**ɔ** 21-t**ɤ** wa̲n35-ɕjɑw51 →
open my joke
E: k^haj55 w**ɔ** 21-t**ɤ** w**ə̲** n35-ɕjɑw51
'teasing me'

b. I: i51 ti55 jɛn21 le̲j̲51 →
one drop eye tear
E: i51 ti55 jɛn21 l**i̲**51
'a tear drop'

In (38a), the error looks at first pass like a non-contextual substitution of [ə] for [a]. However, if we consider the source to be one or both of the other mid vowel phones in the preceding word, then this would be a contextual perseveration/substitution of either [ɔ] or [ɤ] for the [a] in [wan]; since the only allophone of the mid vowel phoneme allowed in the environment [w__n] is [ə], the vowel is realized as this variant. Similarly, in (38b), the error appears to be the non-contextual omission of the [e] in [lej], which causes the glide [j] to be realized as its vocalic variant [i]. However, if we assume that the [e] has been deleted in dissimilation from the preceding mid vowel [ɛ], then this becomes a contextual error. Thus our analysis helps account for a large number of the otherwise non-contextual errors in our data in a principled way.

The only remaining question is whether or not we can determine a basic variant of the mid vowel phoneme. Using the logic we applied to the low vowels, the central variant should be [ɤ] since it is the only of these phones that can occur alone as a full syllable and in open syllables with no pre-nuclear glide conditioning factors. However, this analysis would produce a somewhat asymmetrical phoneme inventory: /i, y, u, ɤ, a/ (see Maddieson (1984: Chapter 8)). Lin (1989) posits [ə] as the basic variant because it is the phonetic center of the category, and both R. Cheng (1966) and Wu (1994) also select [ə] as the basic variant. Furthermore, of all the mid vowels, [ə] occurs following the broadest range of onset consonants, although it must always be in a syllable closed by [n]. Choosing [ə] as the basic variant of the phoneme provides us with the more symmetrical phoneme system /i, y, u, ə, a/; furthermore /ə/ as a phoneme is far more common in the world's languages than /ɤ/ (Maddieson (1984: 252, 257)). For all these reasons we agree that this phoneme category is best characterized as underlyingly /ə/.

5. Summary and Discussion

5.1. *The Vowel Phoneme System of Taiwan Mandarin*

In this paper we examined four hypotheses regarding the underlying vowel categories of Mandarin, and then looked at phonetic, distributional, and especially psycholinguistic evidence regarding the system of TM. We found that none of the previous theories could completely account for the findings in our study, although we concur with R. Cheng (1966) and Lin (1989) that the system consists of five underlying vowel categories. Our first difference from all four accounts is that in this dialect of Mandarin, there are not two differing apical vowels [ɿ, ʅ]; in TM there is only a high central unrounded vowel [ɨ], and there is no retroflex vowel. We provided evidence that this vowel [ɨ] is a conditioned variant of /i/, which occurs only after coronal fricatives and affricates: specifically, [i] and [ɨ] cannot substitute for each other in errors, and these two phones alternate with each other systematically in errors involving a change in the contiguous environment. We agree with the analysis of R. Cheng and Wu who group the apical phones with the /i/ phoneme; we disagree with C. Cheng who considers the apical phones to constitute a phoneme in their own right, and with Lin who derives the apical phones by rule. Second, we find no evidence that [y] is derived from a sequence of /ui/, as proposed by Wu, since [y] can freely substitute for both [u] and [i] in vowel substitution errors. Third, we find a great deal of evidence that [ɛ] is a variant of the mid vowel phoneme only, and not an allophone of the low vowel phoneme. The phone [ɛ] alternates with all of the other 5 mid vowel phones when the contiguous environment changes in 20 errors, whereas there is only one error that suggests an affiliation between [ɛ] and the low vowel phoneme. Thus our analysis disagrees with all four theories on the status of this phone. We agree with all the theories that [a, ɑ] are variants of the same phoneme, and concur with all but C. Cheng that [a] is the basic variant. Finally, we presented evidence supporting the hypothesis that glides are derived from underlying high vowel phonemes. Thus the vowel system which we propose is the following:

(39) The vowel system of TM (Wan and Jaeger)
/i/ → [i, ɨ, j]
/y/ → [y, ɥ]
/u/ → [u, w]
/ə/ → [e, ɛ, ə, ɔ, o, ɤ]
/a/ → [a, ɑ]

If indeed this is the correct system for TM, then we should be able to show that only these five vowels (including all their respective allophones) can substitute for each other in errors, and that no two phones which we have classified into the same phoneme can substitute for each other. In Table 4 above we presented the number of errors from our data in which each vowel phone substituted for the other phones. In (40) below we have collapsed these numbers into the vowel phoneme categories which we are arguing to be the underlying system of TM; this figure shows the number of times any of the phones from each phoneme substituted for any of the phones of any other phoneme. These figures account for all of our data (N = 56, including the vowel/glide substitutions); this table shows that there were no cases where two phones which we consider to belong to the same phoneme substituted for each other.

(40)

	y	u	ə	a
i	19	12	2	1
y		7		
u			2	1
ə				12

In Sections 3 and 4.2 above we gave several examples of the high vowels substituting for each other (Examples (11), (14), (16), (17), (18), (19)). Below are examples of substitutions across other phonemes; in (41a) the mid vowel phoneme /ə/ [ɤ] is anticipated and substituted for the high vowel phoneme /u/ [u]; in (41b) the high vowel phoneme /i/ [i] is anticipated and substituted for the low vowel phoneme /a/ [a]; and in (41c) the low vowel phoneme /a/ [ɑ] is perseverated and substituted for the mid vowel phoneme /ə/ [ə].

(41) a. I: ʐ̩u35-kwɔ21 tʂʰan21-tʂʰu̲55 wən51-tʰi35 t**ɤ** xwa51 →
if cause problem

E: ʐ̩u35-kwɔ21 tʂʰan21-tʂʰ**ɤ̲** 55 wən51-tʰi35 t**ɤ** xwa51
'If (the product) caused a problem . . .'

b. I: tʰa̲55 l**i**35-kaj55 xɑw35-tɕjow21 lɤ →
he leave very-long

E: tʰ**i̲**55 l**i**35-kaj55 xɑw35-tɕjow21 lɤ
'He has been gone for a long time'

c. I: wɔ21 k**ɑ**ŋ55-k**ɑ**ŋ55 kə̲n55 ʂej35 ʂwɔ55 →
I just with who say

E: wɔ21 k**ɑ**ŋ55-k**ɑ**ŋ55 k**a̲**n55 ʂej35 ʂwɔ55
'Who did I speak with?'

Thus all of our vowel substitution errors support this proposed 5-vowel system.

5.2. *A Psycholinguistic Model of Phonological Representation and Processing*

We would like to end this paper by exploring a possible psycholinguistic account of the representation of vowels in lexical representations in Taiwan Mandarin, and the processes by which they arrive at their correct surface forms in speech production planning. In order to do so we must first present a partial speech production planning model, which illustrates how the phonological representations of words and their surface manifestations are related in processing. This model has been developed based partly on our research, and partly on the previous models of Fromkin (1973) Garrett (1988), and Levelt (1989).

In this model, the circled categories in the left-hand column indicate the representation of information in long-term memory. This includes the 'Lemma Lexicon', which represents information about the semantic and syntactic properties of lexical items, and the 'Form Lexicon', which contains underlying phonological representations (as well as information regarding morphological and orthographic structure of lexical items). The 'Syntax' component includes morphosyntactic rules or algorithms, and syntactic templates, while the 'Phonology' component includes knowledge about the phonological units, sequences, and alternations allowed in the language. The right-hand column includes processing components (in rectangular boxes), which represent actual work being done in short-term memory while a particular utterance is being compiled for speaking, and output/input representations (in italics) which are essentially on-line representations of the output from one processing component, which is then input into the next component for further processing. The final representation in this model, the 'Phonetic Level Representation' is the representation which will be formatted by stored motor programs for actual speaking.

The portions of this model which are relevant to the present discussion are the representations in the 'Form Lexicon' and 'Phonology', as well as the 'Phonological Processes' component. After lexical items are selected and inserted into the syntactic string, their phonological forms are activated; the result of these processes is a string with all the abstract phonological representations for each word or morpheme concatenated into the correct order, i.e., the 'Positional Level Representation'. This representation is input into the 'Phonological Processes' component. This component acts on the abstract phonological input in various stages, in levels similar to those posited in non-linear (e.g., autosegmental) phonological theories (Wan and Jaeger 1998). However, our error data show that the first thing to occur is any kind of error in which an abstract segment (or segmental node) is

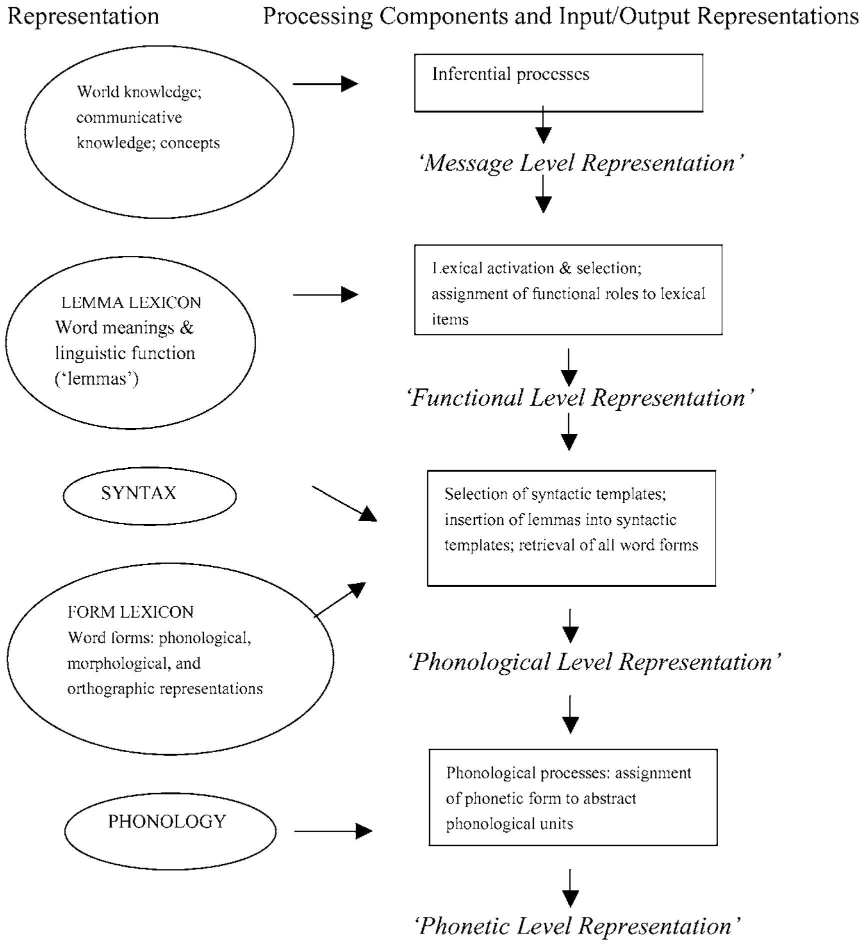


Figure 1. A speech production planning model.

misplaced, added, or omitted in the abstract string. This is because the remaining phonological processes, i.e., both context-dependent rules and feature fill-in redundancy rules which are stored in the 'Phonology' component in long term memory (see below), all operate on the erroneous string *as if it were the intended linear string of phonemes*. Thus the phonological rules of the language 'clean up' any messes made by erroneous placement of segments, by causing the phonemes to be realized in their correct surface forms given the order in which they occur after the errors are made. Let us now put some specific phonological content to this general outline.

5.3. *The Featural Structure of the Vowel System of Mandarin*

We assume that the phonological information in lexical representations is stored in terms of some sort of abstract or underspecified phonological units, similar to the analysis of Lin (1989: 58). In Lin's model, the representation of the vowels of Mandarin can be accounted for in terms of the underspecified feature grid shown in (42).

(42)

	i	y	u	ə	a
High	+	+	+	–	
Low					+
Back	–	–	+		
Round	–	+	+		

Lin also posits the following set of redundancy rules for assignment of the surface phonetic forms for each phoneme:

(43) [ ] → [–high] [ ] → [–round]
[ ] → [+back] [ ] → [–low]

Given our somewhat different analysis of the vowel system of TM, we will need to make some changes in this underlying feature specification in order to account for our findings. First, Lin has specified /i/ as [–back] underlyingly. This was done in her system because the default value for the feature [back] supplied by the redundancy rules is [+back], so the [–back] feature needs to be specified in the underlying representation.[8] However, in our system this would cause us to have a context-dependent rule which would change the feature [–back] to [+back] after the coronal fricative and affricates, in order to derive the surface form [ɨ]. (Note that this designation of the central vowel [ɨ] as [+back] does not mean that it is literally a back vowel; we are using the usual convention of marking central and back vowels [+back] as opposed to front vowels which are [–back].) Thus we propose that in TM, the phoneme /i/ is unspecified for the feature [back]; the [–] value for this feature will be filled in by a context-dependent rule to designate [i], and the [+] feature will be filled in by the redundancy rules to designate [ɨ]. The underlying representation of this phoneme will reflect the fact that the phoneme can be realized as both a [+back] and [–back] variant, which makes the representation more in line with that of /ə/, which is also unspecified for [back] and can have both

[+back] and [–back] surface alternates. Furthermore, a designation on the feature [back] is not necessary in order to distinguish /i/ from all the other vowel phonemes in underlying form, since it is the only category designated as [+high, –round]. Finally, as we will show below, this will allow us to have context sensitive rules which only add feature values, not change feature values.

Second, following the same logic as above, we find that the phoneme /u/ does not need to be specified for the feature [back] (see also note 8). This phoneme is differentiated from all other phonemes without this specification, and the value [+back] will be added by redundancy rules.

Third, our designation of [ɛ] as being part of the mid vowel phoneme solves a problem in Lin's system. Since in her system [ɛ] is an allophone of /a/ as well as /ə/, she needed to have some provision for changing the [+low] specification to [–low] for those surface variants which derive from /a/. In our model all instances of [ɛ] derive from the mid vowel phoneme, and are thus designated only as [–high] in underlying form, and therefore no feature designations need to be changed to derive the surface form. With these three differences in mind, then, we would argue that the vowels of TM can be accounted for in terms of the following underspecified feature grid.[9]

(44)

	i	y	u	ə	a
High	+	+	+	–	
Low					+
Back		–			
Round	–	+	+		

Besides specifying each phoneme so that it is distinct from every other phoneme underlyingly, this underspecified feature grid also captures the fact that the /y/ phoneme is the most marked phoneme in TM, since it has three features specified underlyingly whereas all the other phonemes have only one or two. It is also true that /y/ is a "marked" vowel universally, in the sense that it is much less common in the world's languages than /i, u/, and has the marked combination of [–back, +round] which is disfavored universally (Maddieson (1984: Chapter 8)). In line with it being a marked vowel universally, it is the least frequent of the vowel phonemes in TM, and occurs in the most restricted set of environments, as shown in Table 2 above.

In order to predict the correct surface forms from these various underspecified input forms, we need to have both context-dependent rules and context-independent redundancy rules, as discussed above. These rules will be part of the 'Phonology' in the processing model, that is, the set of rules (or constraints) stored in long term memory which specify what a possible and legal output can be in this language. In (43) above we presented Lin's four redundancy rules, and we agree that these are the correct rules for our new analysis. However, we find that in order to account for all the surface forms of vowels, one further feature is needed. In Wan (1999), the feature [tense] was explored as the possible necessary fifth feature. However, because this feature has little consistent phonetic content (Jaeger (1983), Ladefoged (2001: 80)), it was not found to be adequately explanatory. But a look at the pairs of vowels which need to be distinguished from each other by some feature other than the four listed in Lin's model gives a strong clue as to what this feature actually is phonetically: [e-ɛ], [ə-ɤ], [o-ɔ] and [a-ɑ]. These pairs each represent two vowels with the same basic backness, rounding, and height, but in every case the second of the two vowels is somewhat retracted from the first; in some cases the tongue is pulled directly back, but in others it is pulled back and downward. This analysis of 'retraction' is consistent with formant values of [ɛ] discussed in conjunction with the spectrogram given in (8) above (see note 3). Thus the phonetic property which distinguishes these vowel pairs could be called 'retracted tongue root', or [RTR]; we choose to call this 'retracted' rather than the more common 'advanced', because the retracted versions are in each case the more 'marked' version of the pair (in the sense of less common universally; Maddieson (1984)). The [RTR] feature would be unspecified in all underlying representations; any instance of [+RTR] would be assigned by context-dependent rules, and then the default value of [–RTR] would be assigned by a redundancy rule. Given this feature, a full surface-feature chart of the vowel phones would be as follows:

TABLE 6
Phonetic feature chart, showing specifications of surface vowels

	i	ɨ	y	u	e	ɛ	ə	ɤ	o	ɔ	a	ɑ
High	+	+	+	+	–	–	–	–	–	–	–	–
Low	–	–	–	–	–	–	–	–	–	–	+	+
Back	–	+	–	+	–	–	+	+	+	+	+	+
Round	–	–	+	+	–	–	–	–	+	+	–	–
RTR	–	–	–	–	–	+	–	+	–	+	–	+

In order to derive these fully specified phonetic forms, the following context-dependent rules would be necessary.[10]

(45) Context-Dependent Rules

(a) $\begin{bmatrix}+\text{high}\\-\text{round}\end{bmatrix} \rightarrow [+\text{back}] \;/\; \begin{bmatrix}+\text{coronal}\\+\text{strident}\end{bmatrix} \_\_\_$

$[-\text{back}]$ / elsewhere

(b) $[+\text{low}] \rightarrow [+\text{RTR}] \;/\; \_\_\_ \; [+\text{back}]$

(c) $[-\text{high}] \rightarrow [-\text{back}] \;/\; \_\_\_ \; \begin{bmatrix}+\text{high}\\-\text{back}\end{bmatrix}$

$\rightarrow \begin{bmatrix}-\text{back}\\+\text{RTR}\end{bmatrix} \;/\; \begin{bmatrix}+\text{high}\\-\text{back}\end{bmatrix} \_\_\_ \begin{Bmatrix}[+\text{coronal}]\\ \#\end{Bmatrix}$

$\rightarrow [+\text{round}] \;/\; \_\_\_ \; [+\text{back}]$

$\rightarrow \begin{bmatrix}+\text{round}\\+\text{RTR}\end{bmatrix} \;/\; \begin{bmatrix}+\text{back}\\-\text{cons}\end{bmatrix} \_\_\_ \#$

$\rightarrow [+\text{RTR}] \;/\; \begin{Bmatrix}\#\\ [+\text{cons}]\end{Bmatrix} \_\_\_ \#$

Finally, the redundancy rules which fill in the remaining features are:

(46) $[\;] \rightarrow [-\text{high}]$ $[\;] \rightarrow [-\text{round}]$ $[\;] \rightarrow [-\text{RTR}]$
$[\;] \rightarrow [+\text{back}]$ $[\;] \rightarrow [-\text{low}]$

Combined with the context-dependent rules, these redundancy rules will fill in all of the remaining feature values correctly, so that the 'Phonological Processes' component can output the 'Phonetic Level Representation' in a form which can be used to activate stored motor programs and thus be executed in speaking.

In order to test whether or not we have developed an appropriate model for the underlying representations of the vowel phonemes in TM, let us return to our psycholinguistic data. There are several different factors which come in to play in causing specific vowels to substitute for others in these vowel substitution errors. One factor, mentioned above, is phonological similarity, in that it is well documented that one segment is more likely to substitute for another if the two segments have some phonological properties in common (Nooteboom (1973), van den Broecke and Goldstein (1980), Shattuck-Hufnagel (1986), Jaeger (1992)). Another factor would be the number of environments in which the two vowel phonemes can both occur; if two phonemes have only a few environments in which they contrast, there will be fewer opportunities for the two vowels to substitute for each other. A third factor is frequency; a very infrequent vowel will be involved in substitution errors less often than more frequent vowels by definition.

Looking at the first factor, phonological similarity, it is important to note that the form of the vowel which exists in the 'Positional Level Representation' is the underlying form of the vowel, that is, its underspecified form. This is the representation which is the input to the 'Phonological Processes Component', and thus it is this representation which is active in working memory when phonological substitution errors are made. Thus, we could predict that vowels would be more likely to erroneously substitute for each other if they are more phonologically similar to each other *in their underlying forms*. In order to calculate some measure of the 'underlying similarity' of the vowels, we developed the following metric.

(1) If both vowel phonemes have the same underlying specification for some feature (i.e., both + or both –), award three points; this causes the greatest amount of underlying similarity.
(2) If both vowel phonemes are unspecified for some feature, award two points (since the specification is the same for the two vowels, i.e., unspecified).
(3) If one vowel is specified for a feature, and the other is unspecified for that feature, award one point (since in their surface form they could have the same specification for that feature).
(4) If both vowels are specified for some feature, but have the opposite specification, award no points; this causes the greatest amount of underlying phonological difference.

Using the underspecification feature grid shown in 44 above, we calculated the following 'underlying similarity' measures for each vowel phoneme pair:

(47)

	y	u	ə	a
i	6	7	5	5
y		9	4	4
u			5	5
ə				6

As an example of how these numbers were calculated, consider the underlying featural similarity of [i] and [y]. Both are [+high], giving them a score of 3 points; both are unspecified for low, giving them 2 points; [i] is unspecified for [back] and [y] is specified as [–back], which earns one point; finally, they have the opposite specifications for [round], which earns them no points. Thus their similarity rating adds up to 6.

From these numbers we would make the following predictions: The high vowels would most frequently interact with each other because they are the most phonologically similar; the next most common substitution would be the mid vowel phoneme for the low vowel phoneme; the least most common substitutions would be high vowels for either mid or low vowel phonemes, as they are the least phonologically similar.

Turning to the second factor, we counted the total number of environments in which each of these vowel phonemes can contrast in TM; that is, the number of environments in which each phoneme could substitute for another without any change in environment. The figures are as follows (these figures include the vowel/glide interaction environments).

(48)

	y	u	ə	a
i	11	17	26	27
y		4	4	4
u			15	18
ə				96

Based on these figures, we would expect that the most common error would be low vowels and mid vowels substituting for each other; this would be followed by /i/-low/mid substitutions, /u/-low/mid substitutions, and /i/-/y,u/ substitutions; the least frequent substitutions would be /y/ with /u/, the mid vowel, or the low vowel. It is clear that the ‘phonological similarity’ vs. ‘number of possible contrasting environments’ figures give very different predictions.

As far as the third factor is concerned, we would predict that /y/ in general would be involved in errors less often than any other vowel, since it is the least frequent vowel in TM.

Our actual results show that a combination of factors is at work, but that the phonological similarity factor predominates. The actual number of substitutions for each pair of vowel phonemes was given in (40) above; it is repeated here as (49):

(49)

	y	u	ə	a
i	19	12	2	1
y		7		
u			2	1
ə				12

This chart shows that the most frequent type of error is high vowels substituting for high vowels (N = 38, which includes 13 vowel/glide errors), which is what is predicted by our phonological similarity metric. However, the figure for /u/-/y/ interactions is considerably lower than those for /i-y/ and /i-u/ interactions. This is most likely because of the fact that /y/ and /u/ only contrast in 4 environments, as opposed to 11 for /i-y/ and 17 for /i-u/. The second most common type of substitution involves low and mid vowels, as also predicted by our phonological similarity grid; however, the fact that low and mid vowels can contrast in 96 environments did not disproportionately skew the errors towards this type of error, showing that phonological similarity is overall the more salient feature. Finally, errors involving high vs. mid or low vowels are quite rare; this is predicted by the phonological similarity metric but not necessarily by the 'possible contrasts' metric, since the latter would cause us to expect more errors involving at least /i/ or /u/ and low/mid vowels; the fact that there were no errors involving /y/ and low or mid vowels may be in part due to the small number of possible contrasting environments, as well as the low frequency of /y/. Therefore we conclude that the phonological similarity of vowels in underlying form is the most important factor involved in causing one vowel to substitute for another, although number of possible contrasting environments and frequency can also have some effect. The exact details of the interrelationships among these three factors would have to be worked out in future psycholinguistic research, but it is clear that these three factors are implicated by the pattern of substitution errors found in our data, and that our rank ordering of the importance of these factors is a plausible hypothesis. Furthermore, we would argue that our underspecified feature matrix for the underlying vowel system of TM is supported by these data, because it makes the correct predictions about frequency of interaction of the various vowel phonemes in substitution errors.

5.4 *Conclusions*

Our psycholinguistic data support the following model for the phonological representation and processing of TM vowels. First, vowels are represented in the phonological forms of lexical representations (see Figure 1) in terms of the underspecified feature grid given in (44) above. During speech production planning, lexical items are selected and inserted into syntactic strings, and these abstract representations for the vowels are the form in which vowels occur in the 'Positional Level Representation', i.e., the output of the syntactic processes component. This representation is the input to the 'Phonological Processes' component. Errors involving segments occur when the linear string of phonemes is being input into this processing component; specifically, the substitution, omission, and addition of vowels and consonants (as well as other phonological units) occurs at this point. Then the 'Phonology' is activated from long-term memory, and the abstract units are assigned phonetic form in terms of filling in feature specifications, first by the context-dependent rules listed in (45) above, and then by the general redundancy rules which assign default feature values, as shown in (46). This model completely accounts for all the patterns found in our SOT data, and thus is likely to be a veridical model of the processes and representations involved in producing TM vowels.

As just argued, this psycholinguistic model maps in a simple and straightforward way onto our claims about the underlying and surface forms of TM vowels. Furthermore, the multi-tiered approach of non-linear phonology maps in a simple manner onto our model of the form of underlying representations and of the various processes involved in the 'Phonological Processes' component. The distinction between context-dependent and context-independent (redundancy) rules has a very real instantiation in the psycholinguistic model; it is clear that for the most part the context-dependent rules are specific to the phonology of TM, while the redundancy rules follow universal patterns. Thus our psycholinguistic model can be mapped onto a phonologically plausible theoretical model in a very straightforward way.

Although it is not an absolute requirement that all phonological theories must strive for cognitive reality, this is in fact the stated goal of most phonological theories; that is, they intend to be a 'model of knowledge' which speakers have about the phonology of their languages. However, many theories make such claims without recourse to the kind of psycholinguistic evidence presented in the current paper. We would like to suggest that this type of evidence can be crucial for deciding among competing phonological theories, not only in terms of competing analyses of the phonology of a particular language, as illustrated in this paper, but also in terms of competing types of phonological theories, such as generative theories, non-linear (autosegmental) theories, and optimality theory. Our research suggests that in order for a phonological theory to be cognitively plausible, it must have an explicit theory about the abstract representation of phonological form in lexical representations, some notion of the difference between context-dependent and context-independent processes (which are, after all, what speakers have stored in their 'Phonology', i.e. their knowledge about the phonology of their language), some theory about the distinction between language-specific and universal rules, a set of phonological features which are phonetically motivated, and a plausible theory about the relationships among these various components of phonology. Any phonological theory which not only incorporates all of these theoretical components, but also utilizes both internal and external (psycholinguistic) evidence to support its claims, has the strongest possibility of actually capturing how phonology works in the human mind.

Acknowledgements

We would like to thank Prof. Yen-Hwei Lin and Prof. San Duanmu for helpful comments on this project. Thanks also go to the participants at the LSA Annual Meeting in Chicago in 2000 for valuable suggestions. We also express our appreciation for comments from an Associate Editor of Journal of East Asian Linguistics, and three anonymous reviewers. All remaining errors of analysis or interpretation are our own. This research was supported in part by a National Science Council Grant #892411H004047 to the first author in Taiwan.

Notes

1 Pullyblank (1983) argues that there are no underlying vowels in Mandarin, and Wang (1993) essentially derives all vowels from a feature geometry tree and is not concerned with the underlying phonemic inventory of Mandarin vowels. Thus these analysis were not optimal for examination in this study because of their very different theoretical approaches.

2 When there is more than one consonant in the left-hand column, and one of those consonants is shown in one of the columns under the vowel headings, this is to be taken as shorthand that all of the consonants in the left-hand box can co-occur with this vowel. For example, the entry in the column under [i] which reads [çi] should be taken to indicate that all three palatal obstruents can occur as onsets preceding [i], i.e. [çi], [tçi] and [tçʰi] are all allowable sequences.

3 The fact that the F2 and F3 values of these phones are similar to those of the vowel [ɨ]

shown above suggests that this [ɛ] might be slightly centralized or retracted; a detailed phonetic analysis of the vowels of TM is beyond the scope of the current paper.

4 In this paper we will be using the following notation: C = a non-glide onset consonant; G = a glide; V = the vowel nucleus; F = the final consonant in a syllable, which can be either a glide or a nasal.

5 In Beijing Mandarin, tone 3 is a fall-rise contour, [214]. However, in the dialect under study here, it is a low-fall [21], as also indicated by Yip (1980).

6 In many words which appear as examples in this paper, speakers of Beijing Mandarin would have a retroflex consonant, where speakers of TM have a dental consonant, as for example in [tsɨ55-tsu35] 'content', which would be [tʂɨ55-tsu35] in Beijing Mandarin. In the dialects of many of the speakers in the current study, most retroflex consonants are spoken most often as dentals, particularly in casual speech. The examples given in this paper reflect the exact pronunciations of the speaker who produced these errors.

7 We collected three other errors of the following type: I: tsɨ51-tɕi21 'myself' → E: tɕi51-tɕi21 (meaningless); while these three errors all show an alternation between [ɨ] and [i] when the palatal vs. non-palatal environment is changed, they could also be argued to be cases where one whole syllable substituted for another (but leaving the tone unchanged). Thus we did not include these three errors into our counts of [i]-[ɨ] interactions.

8 Actually Lin's argument is more complicated than this. She argues (1989: 59) that any feature which is involved in co-occurrence constraints on morpheme structure needs to be specified in underlying forms. She cites the following three constraints on vowel sequences: *[+round +round], such as *uau, yau, uəu; *[–back –back], such as *iai, iəi, yai; and *[+high +high], such as *iu, yi, ui; (it is also the case that sequences such as *ji, ɥi, ij, yj and wu, uw do not occur). In our analysis we would have to express the *[–back –back] constraint in terms of a more complex constraint, i.e., *[+high/–round +high/–round]. However, there are no morpheme structure constraints on vowel sequences which require the [+back] feature, so that it does not need to be specified for /u/; Lin makes a similar argument for possibly deleting the [–round] specification in underlying representations, and we follow her line of argumentation.

9 Note that according to some theories of underspecification (Steriade (1995)) we wouldn't need to indicate the [–high] value for /ə/, since it is completely distinguished from the other underlying vowel categories without this designation, and the value of [–] is the default value for the feature [high] which will be assigned by the redundancy rules. However, we need to have this vowel specified as [–high] in its underlying featural grid so that it will trigger the set of context sensitive rules which are applicable to the mid vowel phoneme; since these rules apply before the redundancy rules, this value must be specified in the basic representation. We follow Lin (1989) in including this specification in the underlying feature grid.

10 Although our designation of /i/ as underlyingly unspecified for [back] causes us to assign a feature value to the basic variant by rule, the savings in underlying representation plus the fact that all context-dependent rules add feature values rather than changing them makes this the preferred analysis.

References

Baars, B. J. (ed.) (1992) *Experimental Slips and Human Error: Exploring the Architecture of Volition*, Plenum Press, New York.

Berg, T. (1987) *A Cross-linguistic Comparison of Slips of the Tongue*, Indiana University Linguistics Club.

Chao, Y-R. (1934) "The Non-Uniqueness of Phonemic Solutions of Phonetic Systems," *Bulletin of The Institute of History and Philology*. Academia Sinica 4(4), 363–397.

Chao, Y-R. (1968) *A Grammar of Spoken Chinese*, University of California Press, Berkeley.

Chen, J. Y. (1993) "A Corpus of Speech Errors in Mandarin Chinese and their Classification," *The World of Chinese Language* 69, 26–41.

Chen, J. Y. (1999) "The Representation and Processing of Tone in Mandarin Chinese: Evidence from Slips of the Tongue," *Journal of Applied Psycholinguistics* 20, 289–301.

Cheng, C. C. (1973) *A Synchronic Phonology of Mandarin Chinese*, Mouton, The Hague.

Cheng, R. L. (1966) "Mandarin Phonological Structure," *Journal of Linguistics* 2(2), 135–262.

Cutler, A. (1982) "The Reliability of Speech Error Data," in A. Cutler (ed.), pp. 7–28.

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

Dell, G. (1980) *Phonological and Lexical Encoding in Speech Production*, PhD dissertation, University of Toronto.

Dell, G. (1984) "The Representation of Serial Order in Speech: Evidence from the Repeated Phoneme Effect in Speech Errors," *Journal of Experimental Psychology: Learning, Memory and Cognition* 10, 222–233.

Duanmu, S. (1990) *A Formal Study of Syllable, Tone, Stress, and Domain in Chinese Languages*, PhD dissertation, MIT.

Duanmu, S. (2000) *The Phonology of Standard Mandarin*, Oxford University Press, Oxford.

Fromkin, V. (1973) "The Non-Anomalous Nature of Anomalous Utterances," in V. Fromkin (ed.), pp. 215–242.

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

Fromkin, V. (ed.) (1980) *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand*, Academic Press, New York.

Garrett, M. (1988) "Processes in Language Production," in F. Newmeyer (ed.), *Linguistics: The Cambridge Survey III. Language: Psychological and Biological Aspects*, Cambridge University Press, Cambridge, pp. 69–96.

Goldsmith, J. (1995) "Introduction," in J. Goldsmith (ed.), pp. 1–23.

Goldsmith, J. (1995) (ed.) *The Handbook of Phonological Theory*, Blackwell, Cambridge.

Jaeger, J. (1983) "The Fortis/Lenis Question: Evidence from Zapotec and Jawoñ," *Journal of Phonetics* 11, 177–189.

Jaeger, J. (1992) "Phonetic Features in Young Children's Slips of the Tongue," *Language and Speech* 35, 189–205.

Jakobson, R., G. Fant, and M. Halle (1951) *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*, MIT Press, Cambridge, Massachusetts.

Ladefoged, P. (2001) *A Course in Phonetics* (5th Edition), Harcourt College Publishers, Fort Worth.

Laubstein, A. S. (1987) "Syllable Structure: The Speech Error Evidence," *Canadian Journal of Linguistics* 32, 339–363.

Levelt, W. (1989) *Speaking: From Intention to Articulation*, MIT Press, Cambridge, Massachusetts.

Li, W. C. (1999) *A Diachronically-Motivated Segmental Phonology of Mandarin Chinese*, Peter Lang, New York.

Lin, Y. H. (1989) *Autosegmental Treatment of Segmental Processes in Chinese Phonology*, PhD dissertation, University of Texas, Austin.

Lin, Y. H. (1992) "Vocalic Underspecification in Two Mandarin Dialects: A Case Against Radical Underspecification," *Journal of East Asian Linguistics* 1(3), 219–253.

Maddieson, I. (1984) *Patterns of Sounds*, Cambridge University Press, Cambridge.

Mowrey, R. and I. MacKay (1990) "Phonological Primatives: Electromyographic Speech Error Evidence," *Journal of the Acoustic Society of America* 88, 1299–1322.

Nooteboom (1973) "The Tongue Slips into Patterns," in. V. Fromkin (ed.), pp. 144–156.

Ohala, J. J. (1986) "Consumer's Guide to Evidence in Phonology," *Phonology Yearbook* 3, 3–26.

Pulleyblank, D. (1983) "Vowelless Chinese? An application of the three tiered theory of syllable structure to Pekingese," Paper presented at the XVI International Conference on Sino-Tibetan Languages and Linguistics, at the University of Washington.

Shattuck-Hufnagel, S. (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. Klatt (1979) "The Limited Use of Distinctive Features and Markedness in Speech Production," *Journal of Verbal Learning and Verbal Behavior* 13, 41–55.

Shen, J. (1993) "Slips of the Tongue and the Syllable Structure of Mandarin Chinese," in S-C. Yau (ed.), *Essays on the Chinese Language by Contemporary Chinese Scholars*, Centre de Recherches Linguistiques sur l'Asie Orientale-Ecole des Hautes Etudes en Sciences Sociales, Paris, pp. 139–162.

Stemberger, J. (1983) *Speech Errors and Theoretical Phonology: A Review*, Indiana University Linguistics Club.

Steriade, D. (1995) "Underspecification and Markedness," in J. Goldsmith (ed.), pp. 114–174.

van den Broecke, M. P. R., and L. Goldstein (1980) "Consonant Features in Speech Errors," in V. Fromkin (1980) (ed.), pp. 47–65.

Wan, I. P. (1997) "The Status of Prenuclear Glides in Mandarin Chinese: Evidence from Speech Errors," *Chicago Linguistics Society* 33, 417–428.

Wan, I. P. (1999) *Mandarin Phonology: Evidence from Speech Errors*, PhD dissertation, University at Buffalo, State University of New York.

Wan, I. P. (2000) "Testing the Psychological Reality of Coronals," *Proceedings of the 2000 NCCU Teachers' Conference on Linguistics Research*, National Chengchi University, Taipei, pp. 66–93.

Wan, I. P. (2001) "On the Representation and Processing of Prenuclear Glides in Mandarin," *Proceedings of the 7th International and the 19th National Conference on Chinese Phonology*, National Chengchi University, Taipei, pp. 232–248.

Wan, I. P. (2002) "Feature Theories from a Cognitive Science Perspective," *Proceedings of the 1st Cognitive Linguistics Conference*, National Chengchi University, Taipei, pp. 172–198.

Wan, I. P. and J. Jaeger (1998) "Speech Errors and the Representation of Tone in Mandarin Chinese," *Phonology* 15, 417–461.

Wang, J. Z. (1993) *The Geometry of Segmental Features in Beijing Mandarin*, PhD dissertation, University of Delaware.

Wells-Jensen, S. (1999) *Cognitive Correlates of Linguistic Complexity: A Cross-Linguistic Comparison of Errors in Speech*, PhD dissertation, University at Buffalo, State University of New York.

Wu, Y. (1994) *Mandarin Segmental Phonology*, PhD dissertation, University of Toronto.

Yang, W. (1997) *Speech errors in Chinese: A Psycholinguistic Study*. PhD dissertation, University of Victoria, BC.

Yip, M. (1980) *The Tonal Phonology of Chinese*, PhD dissertation, MIT.

Received: April 7, 2001
Revised: May 3, 2002

Professor I-Ping Wan
Graduate Institute of Linguistics
National Chengchi University
Taipei, Taiwan 116, R.O.C.
E-mail: ipwan@nccu.edu.tw

Professor Jeri J. Jaeger
Department of Linguistics and
Center for Cognitive Science
609 Baldy Hall
University at Buffalo
State University of New York
Buffalo, NY 14260, U.S.A.
E-mail: jjaeger@acsu.buffalo.edu