Phonology **15** (1998) 417–461. Printed in the United Kingdom
© 1998 Cambridge University Press

*Speech errors and the representation of tone in Mandarin Chinese**

I-Ping Wan
Jeri Jaeger
State University of New York, Buffalo

1 Introduction: tones in phonological theory and processing

One of the major advances in phonological theory during the past twenty years has been the refinement of a theory regarding the representation and behaviour of tones (Goldsmith 1976, 1990, Fromkin 1978, Clements & Goldsmith 1984, Hyman 1986, 1993, Pulleyblank 1986, 1997), particularly for African tone languages and East Asian tone languages (for recent reviews see Odden 1995 and Yip 1995). A general outline of such a theory, using an autosegmental framework, might be something like the following: (a) Tones or tone melodies are represented in underlying phonological representations (UR); in some cases they are linked to specific 'tone-bearing units' (TBUs) such as syllables, moras or vowels on other tiers in UR, and in other cases they are unlinked. (b) Phonological rules will associate tones with correct TBUs according to universal and language-specific principles, including a universal well-formedness condition, such that no tones or TBUs which remain at the end of the derivation are unassociated. Tones may be delinked and omitted or reassociated by phonological rules, which may involve tone spreading, tone sandhi and a number of other phenomena. (c) Tones have a primarily lexical rather than syntagmatic function. (d) Contour tones are represented in UR as a sequence of two or more level tones, and function as tone sequences in tone rules, such that for example an underlying HL sequence may surface as either a falling contour tone on a single TBU or a high–low sequence

* We would like to thank the following people for helpful comments on this project: Colleen Fitzgerald, Cori Grimm, Alissa Melinger, Karin Michelson, John Ohala, Chris Yuen and other members of the Crosslinguistic Slips of the Tongue Research Group of SUNY Buffalo. We also express our appreciation for comments from an associate editor of *Phonology*, and two anonymous reviewers. All remaining errors of analysis or interpretation are our own. This research was supported in part by a Research Assistantship awarded to the first author by the Cognitive Science Center, SUNY Buffalo.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

on two adjacent TBUs. (e) It is assumed that tonal systems in all languages follow the same universal principles, which underlie the theory outlined here.

Most phonological theories consider themselves to be a model of the knowledge that speakers have internalised regarding the phonological representations and rules of their language (Chomsky & Halle 1968, Fromkin 1973, Halle & Clements 1983: 1, Goldsmith 1990: 9, Carr 1993: 98, Mohanan 1995). Although 'models of knowledge' do not necessarily map directly onto models of real-time processing developed from the results of various behavioural studies (experiments, slips of the tongue, acquisition, etc.), there are many researchers who feel that the phonological theories which have the best possibility of having cognitive status are those which are strongly compatible with such models (Derwing 1973, Ohala & Jaeger 1986, Ohala 1986, 1995, Jaeger *et al.* 1996, among others). This stance has recently been echoed by Goldsmith (1995a: 20) in his introductory chapter to *The handbook of phonological theory*: he states that while most of the material in the book intends to be 'limited to matters *internal* to phonological theory', nevertheless he acknowledges that an important issue is how phonological theory is linked to such fields as neurobiology, lexical access, speech production and speech perception. He continues to say that 'it is truer today than ever before that the studies bridging components and methodologies are critical for the testing and refinement of phonological theory'.

One methodology which has a long history of use in this respect is the analysis of slips of the tongue, i.e. one-time production errors made online by speakers. These errors have been taken as important evidence regarding the units and structures involved in the cognitive representation of language, the relationships among these units and structures, and the status of various posited rules of grammar (Fromkin 1973, 1980, Cutler 1982b, Stemberger 1983, Berg 1987, among others). Out of such studies has come evidence for the cognitive reality of such phonological phenomena as features, phonemes (both segment-sized units and allophonic variation), productive phonological and morphological processes (e.g. vowel harmony; Kiraly 1996) and syllables, to name a few. However, the majority of slips research has been done on English and related languages (Dutch and German), and therefore there are few reports in the literature which bear on the issue of the representation and processing of tone (see Gandour 1977 for Thai tone; Moster 1991, Chen 1993, Shen 1993 and Yang 1997 for various issues regarding Mandarin slips; Wan 1996, Wan & Jaeger 1996, 1997 regarding Mandarin tone).

The purpose of the present paper is to present evidence regarding the representation and processing of tone in the Mandarin Chinese dialect spoken natively in Taiwan, by examining a corpus of 788 slips of the tongue collected by the first author. We will focus on the four issues raised above, briefly: What is the status of tone in underlying representations? What is the status of tone during phonological processing? Is tone an organiser of phonological properties of the syntagmatic string? Are

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

Mandarin contour tones best represented underlyingly as a sequence of level tones? After presenting our findings, we will discuss how our processing evidence bears on phonological theories of tone.

The plan of the paper is as follows: in §2, a brief description of Mandarin phonology, including phonemes, syllable structure and tones, will be presented, followed by a discussion of the theoretical issues regarding tone to be explored here. The third section of the paper presents our methodology of data collection and analysis, with samples of the types of errors examined, and an explanation of how various findings are interpreted in terms of theoretical issues. §4 presents our results in detail, and the final section discusses our findings with reference to both the representation and the processing of tone in Mandarin, with implications for phonological theory.

2 Mandarin tone

2.1 Phonemes of Mandarin

In this study we assume the following consonantal inventory of Mandarin.

(1)	labial		dental		retroflex		palatal		velar	
	vl	vd	vl	vd	vl	vd	vl	vd	vl	vd
unaspirated plosive	p		t						k	
aspirated plosive	p^h		t^h						k^h	
fricative	f		s		ʂ	z	ɕ		x	
unaspirated affricate			ts		tʂ		tɕ			
aspirated affricate			ts^h		$tʂ^h$		$tɕ^h$			
nasal		m		n						ŋ
liquid				l						
glide		(w)						(j,ɥ)		(w)

There are a number of co-occurrence restrictions between consonants and vowels which are not the focus of the current paper; specifically, the palatal fricative and affricates only occur preceding high front vowels and glides [i y j ɥ], whereas the dental and retroflex fricatives and affricates as well as the velar stops do not occur in palatal environments. The three glides [j ɥ w] can be consistently derived from the underlying high vowels /i y u/ when these vowels occur adjacent to another vowel in UR, and thus are considered to be derived from vowels in our study; we will present some evidence for this analysis below (see also Wan 1997).

We adopt the analysis of Mandarin vowels found in Lin (1989):

(2)	front	central	back
close	i y		u
mid		ə	
open		a	

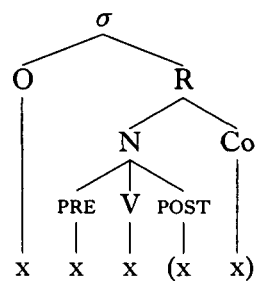
https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

There are complex variations found with vowels; in general, /i/ → [ɪ] after dental and retroflex fricatives and affricates, /a/ → [ɑ] before velars [w ŋ], and /ə/ has the variants [e ɛ ə ɤ ɔ o] under a number of conditions. This analysis will be reflected in our notations of underlying and surface representations in examples presented in the current paper.

In this paper Mandarin examples will be presented in International Phonetic Alphabet symbols, rather than the pinyin romanisation; in some cases we present underlying forms in phonemes between slant brackets, with phonetic realisations in square brackets, both using IPA symbols (as in the preceding paragraphs). We do this for two reasons: first, the actual phonetic characteristics of the utterances under study are crucial to the analysis, and therefore we need to make them systematically explicit. Second, while the use of IPA may be confusing to readers familiar with pinyin, we assume that many readers are not conversant with the pinyin orthography; for these readers, translating from the orthography to the phonetic forms throughout the paper could prove difficult.

The syllable structure of Mandarin is controversial, due largely to the status of the prenuclear and postnuclear glides (Wan 1997). The sequences of glide + vowel or vowel + glide which arise from an underlying sequence of two vowels are not diphthongs, in that the glide retains some phonetic and phonological autonomy. The prenuclear glide has been argued by some researchers to be aligned with the onset, either in a consonant cluster (Bao 1990a) or as a secondary articulation (Duanmu 1990, Wu 1994, Wang 1996); others have argued that it is structurally aligned with the nucleus, but is not a sister with the vowel under the V node, as shown in (3) (Lin 1990). Similarly, the postnuclear glide has been considered to be a constituent with the vowel, making up the nucleus (Chung 1989 for Hakka), while others have considered it a coda consonant (Duanmu 1990, Wu 1994, Wang 1996). The fact that all monophthongal vowels can be followed by a coda nasal, but VG sequences cannot, supports the traditional view that such a sequence is not a diphthong, that is, a structurally unitary constituent. These issues will be relevant to our discussion below of the TBU, as well as the tone features; in this paper we will assume the following syllable structure, as it is the one our slips data support (Wan forthcoming).

(3)



Constraint: a syllable may have 1–4 x's; POST and Co are mutually exclusive.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

2.2 Description of Mandarin tone

Mandarin Chinese has four distinctive tones, and a neutral tone which will be discussed below. The conventionally accepted inventory of tones in Mandarin is given in (4), in terms of the Chao (1930) tone numbers, with '5' indicating the speaker's highest pitch, '3' middle and '1' lowest. Note that these tone numbers are not intended to indicate that the tones are made up of underlying sequences, but simply to indicate relative pitches and pitch contours within each tone (i.e. these are not 'tone features'). Again, we use this system rather than other systems of tone marking because it is optimally transparent regarding the phonetic properties of the tones.

(4) *Mandarin tone inventory*

tone number	*tone*	*example*	
1	high [55]	[ma55]	'mother'
2	rising [35]	[ma35]	'hemp'
3	low-falling [21][1]	[ma21]	'horse'
4	falling [51]	[ma51]	'scold'

Most syllables in Mandarin carry an underlying tone. However, some light (CV) syllables with underlying tone can be produced as unstressed when they are functioning as grammatical particles, as shown in (5a), or as the second element in a reduplicated kinship term (a common alternative form of monomorphemic kinship terms), as shown in (5b). (For 'stress' in Mandarin, see Chen 1984.) In these cases the syllable duration is shortened, and the syllable is spoken with a 'neutral' tone; that is, it loses its underlying tone and is pronounced with a mid-level pitch on the surface. (Neutral tone is unmarked in our examples.)

(5) a. /iŋ21 tsi21/ → [iŋ21-tsɨ] 'shadow'
shadow son shadow-NOUN SUFFIX

b. /ti51 ti51/ → [ti51-ti] 'younger brother'
younger.brother younger.brother younger-brother

Furthermore, there are a few grammatical particles which are never produced with stress, and have no underlying tone; they are always produced with neutral tone, as illustrated in (6):

(6) a. [ni21 jɑw51 tʂʰi55 fan51 ma] 'Do you want to eat?'
you want eat rice Q

b. [ta55 laj35-lɤ] 'He has come.'
he come-PERFECTIVE ASPECT

Although in most cases tones are produced phonetically according to the underlying lexical tone, there are a few tone-sandhi processes in Mandarin. The most prominent case occurs when there are two syllables

[1] In Beijing Mandarin, tone 3 is a fall–rise contour, [214]. However, in the dialect under examination here, it is a low–fall [21], as documented also by Yip (1980).

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

in sequence, both of which have underlying tone 3 [21]; in this case the first syllable changes to tone 2 [35], as illustrated in (7):

(7) /xən21 xau21/ → [xən35 xɑw21] 'very good'
very good

If a syllable has tone 3 underlyingly, but is produced as neutral tone on the surface, it cannot trigger tone sandhi, as shown in (8).

(8) /t^ha55 tɕiə21 tɕiə21/ → [t^ha55 tɕjɛ21 tɕjɛ] 'his older sister'
he older.sister older.sister *[t^ha55 tɕjɛ35 tɕjɛ21]

2.3 Theoretical issues

2.3.1 *The relationship between tones and segments in underlying and surface representation.* Most early generative theories (e.g. Halle & Stevens 1971) regarded tone as a feature of vowels, e.g. [+high tone], [+falling tone] and so on. However, Goldsmith (1976, 1979, 1990), looking predominantly at African languages, found evidence that tone was not as tightly linked to vowels as other segmental features such as [back], and thus proposed Autosegmental Phonology to deal with the more autonomous status of tone. Specifically, he found that tones could shift, spread and be deleted without affecting the quality of vowels or syllables to which they were linked. Evidence from segmentless tones, toneless segments and rules that affect tones or segments separately led Williams (1976) and Yip (1980) also to support the claim that tones should be treated phonologically as suprasegmental, which involves tones being independent from segmental representations at some phonological level. Leben (1978) argued that in underlying representation, tones are on an independent level, but that they are realised as phonetic features on vowels in phonetic representation. For example, in Mende (Mali, Senegal) every word has one of five tonal melodies, which surface with different phonetic representations depending on the number of syllables in the word:

(9) *Mende* (Leben 1973, 1978, Katamba 1989: 195)

	underlying representation	*phonological representation*	*phonetic form*	
a.	HL /mbu/	H L | [mbu]	[mbû]	'owl'
b.	HL /ngila/	H L | | [ngi.la]	[ngílà]	'dog'
c.	HL /felama/	H L | | [fe.la.ma]	[félàmà]	'junction'

In Autosegmental Phonology in general, tones are analysed as being represented on their own autosegmental tier in underlying representations.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

In some cases tones are inherently linked to vowels, and in others they are 'floating tones', i.e. they are not linked to particular segments but are instead a property of the whole word, as in the Mende example above. Phonological rules link tones to appropriate TBUs to produce the final phonetic form. Optimality Theory provides a similar analysis, with 'constraints' such as LINK and ALIGN performing the same work as autosegmental 'rules'. For example, the constraints INTERPRETABILITY[tone] and LINK[tone] constitute two-thirds of Goldsmith's (1976) Universal Association Convention (Pulleyblank 1997).

Regarding Mandarin tone, various researchers have made different claims about the relationship between tones and segments, suggesting that tones are associated with syllables, rhymes, moras or vowels. Wang (1967) and Cheng (1973) argued that the tone unit is associated with syllables in underlying representation, based on evidence from some Chinese languages in which tone features are sometimes linked with the initial consonant, sometimes with the nuclear vowel and sometimes with the final consonant, due to various phonological rules. However even in phonetic form, tone was seen as independent from segments; this claim was based on the fact that segmental features are usually not relevant to the various types of tone sandhi.

Bao (1990a) argued, based on data from *Fanqie* languages (language games) in Mandarin, that tones are associated with rhymes. For example, in the 'mey-ka' pattern, where single-syllable words are mapped into the template [(Onset)ej51 k(Rhyme)], the second syllable in the output word always carries the tone of the original rhyme: /pən21/ → [pej51-kən21] 'book'. Following the framework of feature geometry, Bao (1990b) assumed that at the level of representation prior to the phonetic level, tonal elements are represented as autosegments independent from the segments on which they are realised, but that tone is phonetically realised on the segments which serve as syllabic nuclei.

Woo (1969) claimed that Mandarin tone is linked to the vowel nucleus in underlying representation, and that each tone unit (toneme) is linked to a mora of timing duration. This claim was based on a comparison of the duration of vowels spoken on tone 3, which is analysed as having three tonemes [214], to the duration of vowels spoken on the other tones, which have only two tonemes each. She found that syllables spoken in isolation involving the third tone are phonetically longer than syllables involving one of the other tones of the language. She argued that the syllables with three tonemes also contain three moras of timing, and thus that the TBUs are moras. Similarly, Duanmu (1990) argued that tone features are linked to single mora-carrying segments (single vowels), which predicts that contour tones can only occur on bimoraic (long) vowels (1990: 101); note that this caused him to analyse nearly all vowels in Mandarin as bimoraic.

While it appears that these various theories make different claims about the relationship between tone and segmentals, most of them are in fact compatible with each other. The apparent differences among them may be due to the fact that the various theories focus on representations of

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

segments and tones, and the relationship between the two, at different points in the derivation. A composite theory would be something like that outlined in the introduction above: in lexical representations, tones are autosegmental and are inherently stored with the entire phonological form of the word, either in an unlinked form, or specifically linked with vowels or other aspects of the phonological form. During the phonological derivation of the forms, however, they may become delinked and move, spread, be added or deleted differentially from the segments of the word. Finally, the configuration of tones which has resulted after all phonological processes have applied must be relinked with segments, so that tones can be phonetically realised on voiced segments in appropriate syllables, typically on vowels or other sonorants which can carry pitch. This composite sketch is the theoretical model which we will be testing against the empirical data drawn from our slips of the tongue corpus in later sections of this paper.

2.3.2 *Tone* vs. *stress*. One issue which is of particular interest in this study is the role of stress compared to tone during the phonological organisation of the syntagmatic string. In English and other Germanic languages, it is clear that both lexical stress and overall phrasal intonation play important roles in organising the phonological phrase. Evidence for this role comes from phonological speech errors, where two syllables which are involved in an error (e.g. some segments are exchanged between the two syllables) nearly always have the same stress value, usually both strong, but sometimes both weak; it is rare for phonological elements from a strong syllable to interact with those from a weak syllable in an error (Fromkin 1973: 231, Shattuck-Hufnagel 1983: 125, Berg 1987: 10). Furthermore, the tonic syllable (the most strongly stressed syllable in a phrase, due to its being in the lexical item which is the semantic or pragmatic focus of the phrase) is the most prominent syllable during assignment of phonetic form. In anticipations it is nearly always the source of the erroneous phonological element(s), and in perseverations the tonic word nearly always contains the error; in exchanges it is one of the two syllables involved, and it is almost never deleted. Mandarin does not have a 'tonic syllable' *per se*, as it does not have the same type of intonation system as Germanic languages. However, it will be interesting to see whether tone has the same syntagmatic organising status as lexical stress in phrases, that is, whether two syllables are more likely to be involved in a phonological error if they carry the same tone. Beckman (1986) argues that accent and tone are fundamentally different from each other, based on evidence regarding their different functions, physical properties and speakers' different attitudes toward these two types of phonological systems. Specifically, her evidence suggests that tone is linked more closely with lexical items ('meaning'), while stress is linked more closely with phrasal prosody. Thus we predict that tone will not show the same phrasal organising properties as stress.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

2.3.3 *Are contour tones unitary or sequences of level tones?* Another issue which has been at the centre of an ongoing theoretical controversy for many years is whether contour tones are unitary or sequences of level tones at some level of representation. Early theories considered contour tones to be phonological units with the same status as level tones (Pike 1948, Wang 1967, Cheng 1973), as illustrated in (10a) below for rising tone. However, more recent theories have argued that all contour tones are underlyingly a sequence of level tones. In some models the level tones are considered to be directly linked to TBUs, as shown in (10b) (Goldsmith 1990, Duanmu 1990, 1994).

(10) *Rising tone*

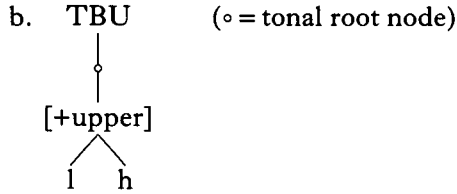


Yip (1980, 1989, 1995) has suggested a different analysis of contour tones, which involves two types of tone features. The first is a register feature [upper], which situates the beginning of the tone in either the upper or the lower part of the speaker's register. The second is the actual tone features 'high' and 'low', which indicate pitch within each register range. (Note that these tone 'features' are different from segmental features such as [voice] in that they are sequential rather than simultaneous; we will return to this point in §4.3.3 below.) Following the discussion in Yip (1995: 478) for Cantonese, we hypothesise that the tones of Mandarin would be represented as in (11a); a diagram of tone 2 (rising) is shown in (11b), where the tone features are linked to a tonal root node, which is itself linked in a unitary fashion to the TBU node.[2]

(11) a.

tone	*pitch*	*register + tone features*
1	55	[+upper, hh]
2	35	[+upper, lh]
3	21	[−upper, hl]
4	51	[+upper, hl]

b.



There are two basic reasons for the change in how contour tones were viewed in linguistic theory, i.e. from being unitary phenomena to being sequences of level tones. First, as mentioned above, the theoretical work

[2] Duanmu (1990) proposes a model of tone similar to Yip's, with 'Register' being captured by the features [±stiff VCs], [±slack VCs] (Halle & Stevens 1971), which give the registers High, Mid and Low; Yip's [h, l] are represented by the features [above] and [below].

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

in autosegmental phonology from which the 'sequences of level tones' model was developed was based primarily on the study of African languages, in which there is a great deal of language-internal evidence in support of this analysis. For example, in the Mende data discussed above (9), the analysis that contour tones on single-syllable words derive from underlying sequences of pitches allows for a very simple statement of possible underlying pitch patterns in all lexical items in the language. Similarly, in Gwari (Nigeria; Hyman 1973, Katamba 1989), there is much evidence that phonetic contour tones are derived from the spreading of underlying tone features.

(12) *Gwari*

underlying representation	*phonological representation*	
[òkpá]	[òkpă]	'length'
L H	L H	
| |	⌈⌉	
/okpa/	[okpa]	

Furthermore, there are a number of cases in African languages where only one feature of a contour tone will project to phonological rules, for example in Luganda (Uganda), where a [hl] falling tone acts the same as a simple low tone [l] in triggering downstep on a following high tone (Katamba 1989: 192). Finally, there are processes involving morphologically complex words that are best explained in terms of tone features; for example, in Mende, adding a toneless suffix can cause a surface contour tone to decompose into its sequence of tone features, allowing the rightmost tone feature to spread to the suffix: *mbû* 'owl' → *mbú-mà* 'on the owl' (Leben 1978). Thus there is ample evidence for the correctness of this analysis for African languages.[3]

The second reason has to do with a theoretical stance common to all phonological theories, which is that the theory should be as general as possible; that is, it should present a unified statement of the representation of phonological phenomena both within individual languages and cross-linguistically. Thus if it can be shown that this particular analysis of tone has explanatory status in some languages, it should be adopted as a universal theory about tone (Yip 1980). However, there have been ongoing indications in the literature that there are intrinsic differences between African and East Asian tone systems (Sapir 1921: 79–80, Pike 1948; see Yip 1989 for a recent summary). First, and most obviously, African languages typically have register tone systems, with two or three distinctive levels of pitch. In these languages, tone patterns are associated with whole phonological words which may be multisyllabic or multimorphemic; tone frequently has grammatical function. Contour tones are

[3] Other non-African tone languages also support this analysis. For example, in Soylatepec Mazatec, only one feature of a contour tone is deleted by tone sandhi rules (Goldsmith 1990: 42).

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

usually derived from phonological processes, and can occur only on margin syllables in the phonological word. On the other hand, East Asian languages typically have contour tone systems, with four to six different tones involving distinctions in both pitch range and direction + slope of the contour. Tones are associated with individual syllables, and any tone can occur on any syllable in a word (Yip 1995). Tone has a predominately lexical function. Secondly, phonological processes involving tones show differences between the two types of systems. For example, African languages typically show downstep, and East Asian languages typically show downdrift. Furthermore, Yip (1995) documents a number of cases in East Asian languages involving reduplication or affixation where there is clearly whole-tone spreading or copying, whereas Odden's (1995) survey of African tone finds only one clear case of a language in which a contour tone projects as a unit rather than a sequence to phonological rules (Grebo, Liberia; Newman 1986). Similarly, Duanmu (1990: 181) finds that Mandarin shows little tone spreading, even compared to other Chinese languages. Third, behavioural evidence from language games provides psycholinguistic evidence for differences between the two types of languages. Hombert (1986), looking at word games in which segmental units are exchanged or substituted, found that in African languages tones typically remained linked to specific syllable positions in the phonological word, whereas in some East Asian languages the tone moved with the segmentals. Similarly, in the Chinese *Fanqie* languages discussed by Bao (1990a; see §2.3.1 above), when rhymes and onsets are separated and inserted into the game template, tones remain linked as whole units to the rhymes. Finally, language acquisition studies find that children learning Mandarin produce tones correctly from the onset of word production (about 12 months; Erbaugh 1992), whereas children learning Sesotho (Lesotho) are still sorting out the complex grammatically triggered tone sandhi rules at age 5 (Demuth 1992). Because of these many differences between the two types of systems, it is not unreasonable to suggest that the systems may require different models of underlying representation, even though this causes complexities for phonological theory. For this reason our data pertaining to the unitary *vs.* non-unitary nature of tones in Mandarin will provide particularly important evidence for assessing the cognitive status of current phonological theories.

3 The study

3.1 Methodology and subjects

The current study is based on 788 slips of the tongue selected from the first author's corpus for their relevance to issues involving tone (Wan forthcoming). These errors were collected between 1995 and 1998 in naturalistic settings from native speakers of the dialect of Mandarin spoken in Taiwan; there were approximately twenty subjects who contributed data. Most errors were gathered from tape-recorded free conver-

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

sations, 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. However, when slips occurred in situations where speakers were not being recorded, the errors were immediately written in a notebook; for each error the author recorded the complete utterance including self-corrections, and relevant contextual information; portions were written in phonetic transcription, when appropriate. Subjects ranged from monolingual to trilingual; all but one spoke Mandarin as their first language, with English and Taiwanese as their other language(s).[4] However, all the errors were collected when the speakers were conversing in Mandarin; there are only eight errors in the corpus (not discussed in this paper) which show a bilingual influence. Therefore we take the majority of these errors to accurately reflect the processing involved in speaking Mandarin.

The usual methodology for collecting slips of the tongue is to rely on the native-speaker linguist's intuitions as to the categorisation of the units in the spoken utterance, that is, what categories in the native language were heard by the native listener. This methodology is subject to some problems of listener bias (Cutler 1982a); furthermore, it has been shown that in laboratory-induced slips, some differences between erroneously produced and intentionally produced consonants can be detected using electromyography (Mowrey & MacKay 1990). One could argue that the perception by a native speaker that a syllable was spoken with a particular tone is a more valid psycholinguistic measure than the actual phonetic properties of the utterance. Nevertheless, we subjected several instances of erroneously produced tones and the same tones produced intentionally in the same tonal environment to fundamental frequency analysis, and found no significant differences between the two (Wan forthcoming). We thus assume that the data to be discussed below are sufficiently reliable to support our analyses.

3.2 Tones in processing

3.2.1 *Processing model.* It is important for our analysis that we adopt a model of speech production planning. This is because speech errors do not map directly onto linguistic theories; speech errors occur during speech production planning and execution, and thus can best be explicated in terms of such a model, that is, during which component or stage of the planning process they occur. However, some speech production planning models map transparently onto linguistic theories, and if we find a fairly close match between such a model and aspects of phonological theory, we can extrapolate from this mapping to a supportable interpretation of specific kinds of speech errors in terms of their implications for linguistic theory.

The phonological theory of tones sketched in §2.3.1 above is

[4] One speaker in the study spoke Cantonese as his first language and Mandarin as his second.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

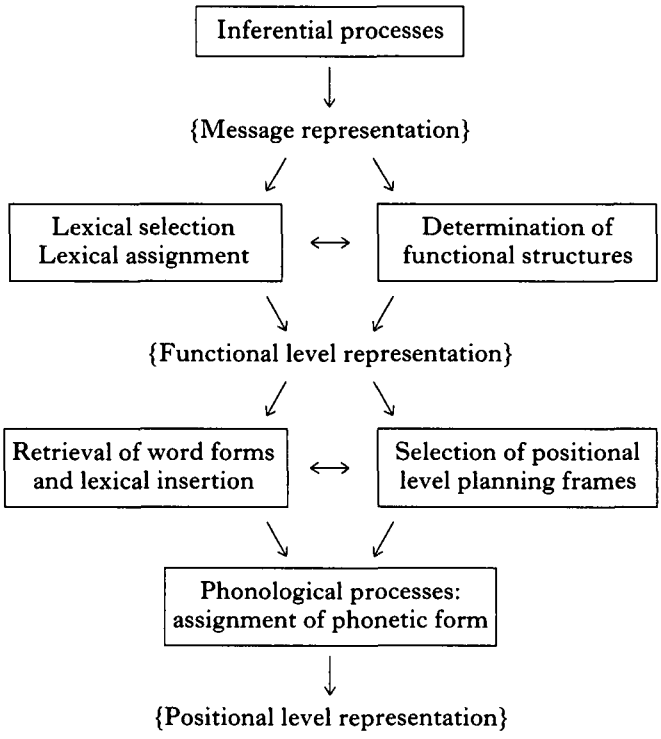


Figure 1
Model of speech production planning (based on Garrett 1984)

'derivational' in the sense that is assumes a sequence of stages or components which take input, perform some operation(s) on it and output it into the next component, deriving the final output in incremental stages. The structure of this theory is very similar to the structure of many of the most thoroughly researched and supported models of speech production planning, for example Fromkin (1973), Dell & Reich (1981), Levelt (1989) and Caplan (1992); most of these models rely heavily on speech error evidence as their primary data. The model to be adopted here is shown in Fig. 1. It is based on the classic model of Garrett (1975, 1980, 1984, 1988), but incorporates some of the refinements made by Bock & Levelt (1994), specifically the double-headed arrows between components which operate in parallel. This model maps particularly well onto autosegmental theory, and has the advantages of showing both the serial and parallel processing evident from speech error data. Although the purpose of introducing this model is to give us a format for making explicit how specific kinds of speech errors are to be interpreted theoretically, the fact that this model works so smoothly in relation to the phonological theory being discussed is implicit support for the cognitive status of both the processing model and the phonological theory.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

In this model there are various processing components which perform increments of the planning task, and produce specific levels of representation in sequence; the last level is the phonetic representation of the utterance. The first component is 'Inferential processes', where the general content of the utterance is selected; the output of this component is the 'Message representation'. The message representation is fed into the 'Lexical selection' component, where candidate lexical items which are semantically appropriate for the message are activated. It is simultaneously fed into the 'Determination of functional structures' component, which, combined with information from the lexical items receiving the highest activation, selects the appropriate syntactic/thematic frames for the utterance. Lexical items are assigned to functional slots in these frames ('Lexical assignment'), and the resultant representation is the 'Functional level representation', an unordered string of lexical items bracketed for syntactic category and function. This representation then undergoes a linear ordering based on the 'Selection of positional level planning frames' component; furthermore, the phonological representations of individual lexical items are retrieved and replace the more abstract lexical representations in the linear string ('Lexical insertion'). At this point the linearly ordered string of phonological representations is inputted into the 'Phonological processes' component. Here phonological processes occur which assign any segmental and prosodic structure not available in underlying phonological representations to the forms, as well as any phrasal prosody; all productive phonological 'rules' take place in this component. The output of this component is the 'Positional level representation', which is a fully specified phonetic string.

3.2.2 *What can go wrong.* Errors can occur at any stage of the speech production planning process. Those which are of interest in this paper occur mainly during lexical selection/insertion and assignment of phonetic form; examples and numbers of each type of error in our corpus are shown in Table I. Terminology is explicated in (13).

(13) A. *Phonological errors*

i. *Directionality of error* (the linear relationship between the error and source)

a. Perseveration: the error occurs after the source.
b. Anticipation: the error occurs before the source.
c. Anticipation/Perseveration (A/P): a potential source occurs both before and after the error.
d. Exchange: two units exchange their positions.
e. Non-contextual: there is no obvious source in the utterance context.

ii. *Type of error*

a. Substitution: one unit is substituted for another (i.e. the target unit is deleted and replaced by the source unit).

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

b. Addition: one unit is added in a new location in the utterance.
c. Omission: one unit is omitted; either the resulting form is more similar to a neighbouring unit (assimilatory), or the error drops a unit which occurs in neighbouring units (dissimilatory).
d. Telescoping: the syntagmatic string collapses into a shorter utterance with at least one rhythmic beat (i.e. vowel) omitted.
e. Multiple errors: combination of types of errors as illustrated above.

B. *Lexical errors*

a. Word blend: two lexical items are activated, and are phonologically blended together and inserted into the same syntagmatic slot.
b. Lexical substitution: one lexical item is substituted for another, and it is clearly a lexical rather than phonological error.
c. Lexical omissions: a lexical item in the utterance is omitted.
d. Misordering: anticipations, perseverations or exchanges of whole words; both target and interfering units are in the discourse context.

While the most appropriate lexical items are being activated by the message representation, lexical items which are semantically related to them will also be activated; furthermore, when a particular lexical item has been selected and its phonological form retrieved, words which are phonologically related to it will have activation spread to them. When an error of 'lexical substitution' occurs (see example A-1 in Table I), it is because one of the semantically or phonologically (or both) similar words mistakenly receives higher activation than the intended word, so is selected and inserted into the functional/syntactic string. A 'lexical blend' error (A-2) occurs when two words which are both appropriate for the meaning and functional representation of a particular slot are equally activated; in this case the system fails to uniquely select one or the other and associates both with a single functional slot. Thus when their phonological representations are activated, it is a phonological blend of the two which is assigned phonetic form and uttered.

Syntagmatic lexical errors can occur when a lexical item is inadvertently omitted (B-1); this may occur because it has not been associated with a functional/syntactic slot, or the wrong functional/syntactic frame has been selected. Another type of error occurs when lexical items selected for the utterance are inserted into the incorrect syntactic slots (misplacement, B-2).

Phonological errors occur after phonological representations have been activated and while phonological processes are assigning the phonetic form to the utterance. They are usually 'contextual', in the sense that the

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

A. *Paradigmatic lexical errors*

1. *Lexical substitutions* (N = 133)

   I: mej35 jow21 i35-kɤ̆ kʰoŋ55-tɤ̆ → E: mej35 jow21 i35-kɤ̆ **man21**-tɤ̆
   not have one-CL empty — not have one-CL full
   'There was no empty (bottle)' → 'full (bottle)'

2. *Lexical blends* (N = 20)

   I: wən51 tsʰən35/laj51 → E: wən51 **tsʰaj51** 'Ask Chen/Lai'
   ask Chen/Lai

B. *Syntagmatic lexical errors*

1. *Lexical omissions* (N = 12)

   I: ni21 tʂɤ̆51 kʰu51-tsɨ jow21 kow51 ta51 →
   you this pants have enough big
   E: ni21 tʂɤ̆51 kʰu51-tsɨ jow21 **Ø** ta51
   'your pants are very big' → (ungrammatical)

2. *Lexical misplacements* (N = 25; includes anticipations, perseverations and exchanges; substitutions and additions)

   I: **ny21** pu51 **ny21**, nan35 pu51 nan35 →
   female not female, male not male
   E: ny21 pu51 ny21, nan35 pu51 **ny21**
   'Girls don't act like girls, boys don't act like boys.' → '...like girls'

C. *Phonological errors*

1. *Features* (N = 44)

   I: tɕin51 **tʂʰ**ɤ̆55 → tɕʰin51 tʂʰɤ̆55
   in car
   'get into the car' → (ungrammatical)

2. *Segmentals* (consonant, vowel, cluster, rhyme, syllable; perseveration, anticipation, exchange; substitution, addition, consonant omission; N = 441)

   I: tʰɑŋ55-mu21 kʰɤ̆51-l**u**21-sɨ55 → E: tʰɑŋ55-mu21 kʰɤ̆51-l**u**21-s**u**55
   Tom Cruise
   'Tom Cruise' → 'Tom (?)'
   (vowel [u] is perseverated from /lu21/, substituted for /ɨ/ in /sɨ55/)

3. *Telescoping* (vowel, rhyme or syllable omission; N = 30)

   I: kən55 **tø**jɛ35 **i55**-mwɔ35-i35-jɑŋ51 → E: kən55 **tøi55** mwɔ35-i35-jɑŋ51
   with sister one-size-one-kind
   'exactly the same as my sister' → (ungrammatical)

4. *Tone errors* (anticipation, perseveration, exchange; substitution or addition; N = 83)

   I: ni**21** laj35 kan51-ma35 → E: ni**21** laj35 kan**21**-ma35
   you come why — you come catch (VB)
   'Why did you come here?' → (ungrammatical)

[Table I. Types of errors analysed in the study (N = 788). 'I' = intended utterance; 'E' = erroneous production. Target (intended utterance) is underlined; **source** (source of interference) is boldfaced; and **error** is underlined and boldfaced.]

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

'source' of the error occurs in the utterance, and causes the 'target' (intended utterance) to be produced erroneously. (In the examples below, the 'source' is shown in boldface, the 'target' is underlined, and the 'error' is boldfaced and underlined.) The error is rarely 'non-contextual', in the sense that there is no source in the utterance. The source can occur after the target (anticipation), or before the target (perseveration), or two units can be exchanged. Errors include substitution of one unit for another, addition of a unit in a new location, omission of a unit or exchange of two units, and can involve features (C-1), single consonants, consonant clusters, single vowels (C-2), rhymes and whole syllables. A special class of phonological error is 'telescoping', in which timing units (vowels, rhymes, syllables) are deleted and the remaining phonological string is collapsed into a shorter string (C-3). Of particular interest in this paper are the errors of tone (C-4), in which tones are substituted, added or exchanged.

In making a decision about the classification of a particular error, four principles were used. First, we assumed the 'minimal movement' principle (Laubstein 1987), which means that the simplest or most conservative analysis was always chosen, involving the smallest segmental error unit. For example ('I' = intended utterance; 'E' = erroneous production):

(14) I: **kɑŋ**55- ts^{h}aj35- tʂ <u>tɑw55</u> → E: **kɑŋ**55- ts^{h}aj35- tʂ **<u>k</u>**ɑw55
previous-MOD knife
'the knife that was just here' → (ungrammatical)

In this example, one could argue that the onset, nucleus and tone of [kɑŋ55] were perseverated and substituted for the onset, nucleus and tone of [tɑw55]. However, the simplest analysis is that only the onsets were involved in the error, and that the phonological similarity of the remainder of the two words facilitated the error. This facilitative 'repeated phoneme' effect is well documented for English (Dell 1984).

Secondly, it is well documented that phonological errors frequently produce real lexical items in the language (Fromkin 1973); this is likely to be even more common in Mandarin than in English, due to the fact that most Mandarin words are monosyllabic, and a change in one or more segments nearly always results in another real word in Mandarin. However, phonological errors can be distinguished from true paradigmatic lexical substitution errors in that lexical errors nearly always preserve lexical category, and are usually semantically related to the intended word; a phonological relationship is less common. Phonological errors, on the other hand, frequently violate lexical category and have no semantic relationship to the target word. For example:

(15) I: wɔ21 ʂi51 <u>lɑw</u>21 **ta**51 → E: wɔ21 ʂi51 **<u>ta</u>**21 **ta**51
I am old big I am old hit
'I am the eldest' → (ungrammatical)

In this case the error [ta21] is a verb 'to hit' whereas the target [lɑw21] is

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

an adjective 'old', so the error word is not of the same lexical category as the target, it has no semantic relationship to the target and is not phonetically similar other than having the same tone. This error is analysed as a case where the segments of the syllable [ta] were anticipated and substituted for the segments of the syllable [lɑw21], leaving the tone of [lɑw21] in place. Similarly, the error syllable [kɑw55] in example (14) above can refer to 'cake' in compounds or can be the adjective 'tall', but in the intended utterance it is meaningless and ungrammatical; however, in this case there is more phonetic similarity between the target and source.

Thirdly, an error was classified as 'non-contextual' only when no source identical to the error unit(s) occurred in the utterance. In (16a) there are no other dentals in the utterance, and in (16b) there is no other tone 4 [51] in the utterance:

(16) a. I: wɔ21 pu35 joŋ51-k̲oŋ55 → E: wɔ21 pu35 joŋ51-t̲oŋ55
I not use-hard.work
'I am not diligent' → (ungrammatical; [toŋ55] = 'winter' or 'east')

b. I: tɑw21-jow3̲5̲ xaj35 ʂwɔ55 → E: tɑw21-jow**5̲1̲** xaj35 ʂwɔ55
travel guide also say
'the travel guide also said' → (ungrammatical)

Finally, a segmental error was classified as a feature error only when there was no segment in the utterance identical to the error segment, but a nearby segment contained the erroneously produced feature value.[5] For example, in error C-1 in Table I, there is no segment [tɕʰ] in the utterance, so the error is taken to reflect the anticipation of the positive value on the feature [aspiration] from the segment [tʂʰ], which occurs later in the utterance. In order to constrain possible sources and not make this classification tautologous, a segment was considered to be a possible source of a feature error only if it occurred within two syllables of the target segment, was different from the target segment by only two features (the feature involved in the error and one other) and, in most cases, occurred in the same syllable structure position as the target. For example, in the error in Table I, the segments [tɕ] and [tʂʰ] are one syllable apart, both occur in onsets, and differ in terms of place of articulation and aspiration, but have identical values on all other features. Similarly, the following error was considered an error involving the feature [place]:

(17) I: t̲ʰɑŋ51-tɤ **k**ɑŋ55-xɑw21 → E: **k̲ʰ**ɑŋ51-tɤ **k**ɑŋ55-xɑw21
hot-LINK enough
'hot enough' → (ungrammatical)

In this error there is no /kʰ/ in the utterance which could have been the whole-segment source of the error segment [kʰ] in [kʰɑŋ]. However, there is a velar consonant [k] in the onset of the syllable two syllables to the right of the target word [tʰɑŋ], and therefore we analyse this as a case where the

[5] The analysis of tone featural errors will be discussed in detail below.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

velar place of articulation from /k/ has been anticipated across two syllables and substituted for the alveolar place of articulation of the target onset consonant [t^h], thus producing [k^h]. This analysis is reinforced by the fact that the rhymes following the onset consonants in both target and source syllables are identical, and that similarity of neighbouring segments is a strong factor in predisposing segments to interact in errors (Dell 1984). Therefore, this is an unambiguous feature error.

One could make an argument that many errors which look like whole segment errors are in fact feature errors, i.e. that only the feature value rather than the whole segment has been substituted. For example:

(18) I: xən35 $\underline{p^h}$u21-**t^h**oŋ55 → E: xən35 **$\underline{t^h}$**u21 **t^h**oŋ55
very common
'very common' → (ungrammatical)

In this error one could argue that it was not the entire segment [t^h] which substituted for [p^h], but rather that only the alveolar feature was anticipated (or spread) from [t^h] and/or perseverated from [n], and substituted for the labial feature of [p^h], with the remaining feature values unchanged. However, there are a number of clear reasons why researchers working with speech error data have systematically considered errors of this sort to involve whole segments rather than features. First, in all corpora published or analysed to date, contextual errors make up the majority of phonological errors; that is, in most cases the whole segment spoken erroneously occurs elsewhere in the utterance, usually nearby. If individual features were free to substitute for each other, one would expect many more non-contextual errors (where the whole-error segment does not occur in the utterance) in which a featural analysis was possible, but in fact both pure non-contextual errors and unambiguous feature errors are uncommon. Secondly, two segments usually interact with each other in an error if they are phonetically similar to each other, occur in the same syllable-structure position, and occur in syllables of the same level of stress (in English). This means that segments which interact in errors are almost never contiguous, and usually have at least one other consonant intervening in the syntagmatic string (e.g. a coda consonant). Therefore it would be very difficult to motivate a feature-spreading analysis (either phonetically or phonologically) in which feature values regularly spread from, for example, one onset C to another onset C two syllables distant, across a number of other intervening segments. Thirdly, errors which involve the addition or omission of a whole segment are quite common, and these have no obvious featural explanation.

Shattuck-Hufnagel & Klatt (1979) provide strong evidence against a featural analysis for segmental errors of the type in question. They looked at all the consonant exchange errors in their corpus in which the two consonants differed by at least two features (N = 90), and found that the majority of errors (N = 70) were eligible for single feature exchanges; i.e. if a single feature had exchanged, the result would have been existing English segments in legal syllable-structure positions, even though neither

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

segment would have occurred in the intended utterance. However, there were only three errors in their corpus in which features but not whole segments unambiguously exchanged, which they take as evidence that actual feature errors are rare. They further found that individual consonants were involved in the 1620 segmental substitution and exchange errors in their corpus in proportions which were statistically correlated with the frequency of occurrence of each segment in English speech. Since the frequency of occurrence of individual segments cannot be related in any consistent way to specific phonetic features, Shattuck-Hufnagel & Klatt argue that it is clearly whole segments and not features which are available to the speech production processing mechanism at the time at which these errors are made. They present a model in which segments from underlying phonological representations are selected and inserted into syntagmatic slots by a 'scan-copier' mechanism; if segments are phonetically similar and in similar syllable-structure slots they are eligible for misselection and misinsertion, but features themselves are not available to this mechanism. Fromkin (1973: 223–227) came to a similar conclusion, arguing that features are not 'independently controlled' in speech production, and therefore are not as susceptible to error as whole segments. Therefore we follow a well-documented principle in classifying an error as 'featural' only in the absence of a whole-segment source.

3.3 Questions to be asked regarding errors and tone

Four groups of questions will be asked with reference to the error data, and the errors will be interpreted as follows:

(i) What is the status of tones in underlying phonological representations? Is there any evidence regarding the association of tone with other aspects of the representation (vowel, rhyme, syllable)? When one word is substituted for another, does it keep its lexical tone or can it be affected by the tone of the other word? If a word is deleted, what happens to its tone? When two words are blended into one, does the error form always contain one of the two underlying tones, or can a hybrid or blended tone be produced? If one of the two underlying tones is produced, is it associated with a particular segmental or structural unit in the UR? Are two words more likely to be involved in paradigmatic lexical errors if they have the same underlying tone, suggesting that tone is a component of underlying phonological similarity among lexical items?

(ii) What is the relationship between tones and segments after lexical insertion, during the mapping from phonological representation to phonetic form? If a segmental unit is involved in an error, are tones affected? What happens to tones when timing units (vowels, rhymes, syllables) are omitted in telescoping errors? If a tone is substituted, exchanged or added, does it affect the segments? How do phonological errors affect tone sandhi; does sandhi occur before or after the errors occur?

(iii) Are units from two syllables with the same tone more likely to be involved in a phonological error, suggesting that tone has the same

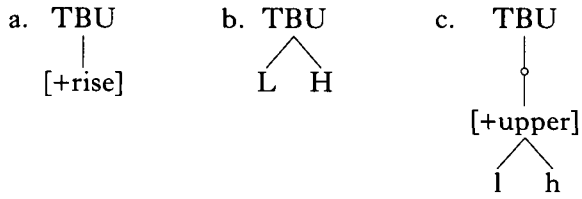
https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

phrasal-organisation function as stress does in languages such as English?

(iv) Are Mandarin contour tones unitary or a sequence of two level tones? If they are sequences underlyingly, we would expect errors in which each tonal component acts as autonomous from the other. This could show up as tone blends in lexical blend or telescoping errors. These would be expected to produce real Mandarin tones, since error productions are filtered through the speaker's native phonology (for example, in English, errors rarely create non-English segments; Fromkin 1973: 229, but see Mowrey & MacKay 1990). The only possible blends which would produce legitimate Mandarin tones distinguishable from both sources would be [55+21]→[51] and [51+35]→[55]. However, if an occasional error actually caused a non-Mandarin hybrid tone to be produced, this would be strong evidence of the autonomy of each tone feature in sequence. We would also expect numerous tone-substitution errors that could be explained as a spreading of tone features from adjacent syllables. This could involve a spreading of tone features from both a preceding and following tone to produce an entirely new tone (e.g. a sequence of [35 21 51]→[35 55 51]), or the spreading of the [upper] feature; the details of possible tone spreadings will be discussed below. On the other hand, if contour tones in Mandarin are unitary and not made up of sequences of tone features, we would expect no errors which require a feature spreading or blending explanation, and would see tones functioning as whole units in every error.

The theoretical implications of possible types of tone errors bear on the models for the representation of contour tones discussed in §2.3.3 above; these are repeated here as (19).

(19) *Rising tone*



A total lack of tone-splitting errors would support (19a), while a preponderance of tone-splitting errors would provide evidence for model (19b). Evidence for the model in (19c) would be found if our analyses of the tone errors showed them to be analogous to segmental errors in terms of the relationship between either whole-segment errors *vs.* feature errors, or unitary segment errors *vs.* complex errors (e.g. affricate errors). We will briefly discuss these two possibilities here; more details will be provided in the results section below.

As discussed in §3.2.2 above, segmental feature errors have been found in previous studies to be much less common than errors involving whole segments, but they do occur in enough cases to warrant the conclusion that features exist as part of the internal structure of segments. Fromkin, for

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

example, presents 265 errors involving single consonant or vowel substitutions or exchanges; of these, 55 (21%) are featural and 210 (79%) are segmental (Fromkin 1973: 243–255). In the case of segments, one could argue that the consonant or vowel root node is the most common unit of error, that is, the unit susceptible to being substituted, added, omitted or exchanged; however, the fact that individual feature values can also occasionally be the error unit argues that features exist as part of segment structure but are less available to the speech production planning mechanism (less 'independently controlled', as Fromkin 1973 puts it). Thus if we find a majority of whole-tone errors, but a reliable subset of split-tone errors or [upper] register errors, we could make the same analysis for tones: that the tonal root node is the more independently controlled unit in speech production planning, but that there is empirical evidence for the existence of underlying tonal features.

On the other hand, errors in which complex segments such as affricates, prenasalised stops or diphthongs are split into component parts are extremely rare, and in some cases claimed to never occur (Stemberger 1983). Since there is ample phonetic evidence that these segments do consist of a sequence of segmental elements, the fact that they rarely split apart in speech errors is strong evidence for the hypothesis that it is segmental root nodes only that are eligible for errors in speech production planning. Thus, while a complete lack of tone-splitting errors could be taken as evidence for the unitary model (19a), it could also be construed as evidence for (19c), since the inability to split apart elements below the root node could simply be a property of the processing of complex phonological units, both segments and tones.

4 Results

4.1 Tones in lexical errors

Because Mandarin tones are lexical, they must be stored in lexical phonological representations along with the segmental representations, and our data are commensurate with this claim. Our data do not provide strong support for the preassociation of tones with any specific aspect of the segmental representation, although taken altogether they suggest that the tone may be linked with the rhyme in lexical representations. The evidence is as follows.

First, when one word is substituted for another in a paradigmatic substitution, the error word always retains its underlying tone. There are 133 such errors in our corpus, and no exceptions; this is true for both content words, as in (20a), and function words, as in (20b).

(20) a. I: jɑw51 swej21 → E: jɑw51 $\underline{\text{t}^{\text{h}}\text{ɑŋ55}}$
want water want soup
'want water' → 'want soup'

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

b. I: tʰa55 ɕin55-tɕʰiŋ35 pu51 xɑw21 →
he mood not good
E: **wɔ21** ɕin55-tɕʰiŋ35 pu51 xɑw21
I mood not good
'He wasn't in a good mood.' → 'I wasn't in a good mood.'

Secondly, if any word or morpheme is added (21a) or misordered (21b), it retains its tone intact; if any word is omitted (21c), its tone is also omitted. We have 37 such errors and no exceptions.

(21) a. I: wɔ21 pu51 y35-kʰwaj51 →
I not happy
E: wɔ21 **kʰwaj51** pu51 y35-**kʰwaj51**
'I am not happy.' → (ungrammatical)
b. I: tɕja55 ʂɨ51 tsaj51 i55-tʂoŋ55 kʰɑw21 →
Jia is in first-senior exam
E: tɕja55 ʂɨ51 tsaj51 **kʰɑw21 i55-tʂoŋ55**
'Jia is taking an exam in the First Senior High.' → 'Jia is taking an exam to get into the First Senior High school.'
c. I: tsaj51 kɤ51-tiŋ55 kʰan51 tjɛn51-ʂɨ51 →
in living room watch TV
E: tsaj51 **Ø** tiŋ55 kʰan51 tjɛn51-ʂɨ51
'watching TV in the living room' → (ungrammatical)

Thirdly, when two lexical items are blended phonologically and inserted into a single syntagmatic slot, the spoken tone is always the tone of one of the two target words; there are no instances of blended tones. Of the 20 lexical blend errors, there were nine cases in which the two monosyllabic target words had the same tone underlyingly. In four of the remaining 11 cases, the blend consisted of the first morpheme from one compound and the second morpheme from the other compound; in three of these errors, both compounds had the same tone pattern, but in the one error in which the two targets had different underlying tones, each spoken morpheme retained its underlying tone:

(22) I: jɑw51 tsɨ55 pʰiŋ35-kwɔ21/ɕjɑŋ-tɕjɑw55 →
want eat apple/banana
E: jɑw51 tsɨ55 **pʰiŋ35-tɕjɑw55**
'I want to eat an apple/banana.' → (ungrammatical)

Four of the remaining cases in which the two target words had different tones provide evidence that the tone may be linked with rhymes in lexical phonological representations. In these cases, the onset from one syllable and the rhyme from the other syllable were phonologically blended; in all cases the tone spoken in the error utterance was that of the retained rhyme, as illustrated in Table I, example A-2, as well as in (23). There were no clear cases involving the onset and tone from one target and the rhyme from the other, suggesting that tones may be linked with rhymes in UR.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

(23) I: wɔ35 ɕjɑŋ21 ʂwɔ55/tʰan35 → E: wɔ35 ɕjɑŋ21 **ʂan35**
I want talk/chat
'I want to talk/chat.' → (ungrammatical)

The remaining three cases are somewhat less clear. At first glance they appear to be cases where the blend involves the segments from one target and the tone from the other, which would suggest that tones are not linked underlyingly and thus either tone may surface in a blend. However, in all three cases, the two targets share the same onset; in one they also share the same underlying vowel (example (24a), where [fej] is underlyingly /fəi/), and in another the vowel from one word surfaces as a glide in the blend (example (24b), where underlyingly /y/ surfaces as [ɥ] in the error; recall that glides are considered to be part of the nucleus, as discussed in §2.1 above).

(24) a. I: ɕɥɛ35 fej51/fən55 → E: ɕɥɛ35 **fej55**
study fee/points
'tuition/credit hours' → (ungrammatical)
b. I: tɕy51 njən35/tɕʰjɛn35 njɛn35 → E: **tɕʰɥɛn51** njɛn35
last year the year before last year
'last year/the year before last year' → (ungrammatical)

These examples are difficult to interpret. Because of the relatively small number of blend errors in our corpus, and the variations in the pattern of tone realisation, we cannot make any strong claims about the status of tone in phonological representations. Taken together with the telescoping data discussed below, the blend data suggest that tones may be linked in UR to rhymes, but much more evidence is needed to fully support this hypothesis.

Finally, tones are clearly part of the phonological organisation of the lexicon. In 133 lexical substitution errors, there are 65 (49 %) cases where the target and source morphemes have the same lexical tone, which is twice as often as would be expected by chance in a four-tone system (if all tones occurred with equal frequency; see §4.2.4 below). These figures partially depend on what kind of morpheme is involved: in the 118 content word substitutions, 54 (46 %) have the same tone; in the six function word substitutions, only two of them have the same tone. There are nine function morpheme substitutions (grammatical particles), and in all nine the target and source tones are identical. Furthermore, in the 20 word-blend errors, in nine monosyllabic errors and three compound errors, the two targets have the same tone (60 %). Taking all 153 paradigmatic lexical errors together (substitutions and blends), we find that in 50·5 % of the errors the two affected lexical items share the same tone, showing that tone is one factor involved in the 'phonological similarity' of lexical items, and thus is a factor in misselection during speech production planning.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

4.2 Tones in phonological processing

After lexical items have been inserted into syntactic frames and their phonological representations activated, phonological processing occurs, which causes phonetic form to be assigned to the string. Our data provide overwhelming evidence that during this processing, segments and tones are unlinked and autosegmental, in that both segments and tones are nearly always involved in errors independently from each other; of 597 segment or tone errors, we have two exceptions.

4.2.1 *Segmental errors.* When consonants or consonant + glide sequences are substituted, exchanged, added, deleted or moved, tones are not affected (N = 308, no exceptions); this is true for both word-initial consonants (see examples (14), (16a), (17), (18) above) and word-final consonants, for example:

(25) I: faŋ51 jɛ**n**21 wɑŋ51 tɕʰy51 → E: fa**n**51 jɛ**n**21 wɑŋ51 tɕʰy51
put eye look go ([n] *anticipated, substituted for* [ŋ])
'take a broad view' → (ungrammatical)

More interestingly, when one vowel is substituted for or exchanged with another, tones are not affected (N = 28, two exceptions; see below), as shown in the following example, where the vowel [ɑ] is perseverated and substituted for [o], but the target word retains its original tone.

(26) I: t**ɑ**w51 ɕj**ɑ**w51-ɥɛn35 ʐ**ɑ**w51 i51-tʂ**o**w55 →
arrive campus around one circle
E: t**ɑ**w51 ɕj**ɑ**w51-ɥɛn35 ʐ**ɑ**w51 i51-tʂ**ɑ**w55 ([ɑ] *perseverated, substituted for* [o])
'go around the campus' → (ungrammatical)

Similarly, when rhymes (27a) or whole syllables (27b) are substituted or exchanged, tones are not affected (N = 104, no exceptions).

(27) a. I: tʰ**o**ŋ35-ɕɥɛ35 t**wɔ**55-pan51 → E: tʰ**o**ŋ35-ɕɥɛ35 t**o**ŋ55-pan51
student most (*rhyme* [oŋ] *perseverated, substituted for* [wɔ])
'most students' → (ungrammatical)
b. I: wɔ21 ʂɨ51 l**ɑw**21 **ta**51 → E: wɔ21 ʂɨ51 **ta**21 **ta**51
I am old big (*syllable* [ta] *anticipated, substituted for* [lɑw])
'I am the eldest.' → (ungrammatical)

The two exceptions to the claim that tones and segments are independent from each other during phonological processing are somewhat ambiguous cases where it appears that one vowel has been substituted for another and has retained its original tone, exemplified in (28) (the other case is analogous to this one).

(28) I: maj21 kʰa**55**-fej55 → E: ma**55** kʰa**55**-fej55
buy coffee
'buy coffee' → (ungrammatical)

The simplest analysis of this error is that the rhyme + tone from [kʰa55] was anticipated and substituted for the rhyme + tone of [maj21]. However,

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

this example could also involve multiple errors: omission of [j] in [maj] in anticipation of the simple rhyme of [k^ha] (assimilation omission), and anticipation of the tone [55], substituting it for tone [21]. In this case the error would not be an exception to the claim of independence. However, the fact remains that the large majority of errors unambiguously support the autosegmental analysis.

4.2.2 *Telescoping.* As explained above, 'telescoping' errors involve the deletion of a rhythmic unit (vowel, rhyme or syllable), and collapsing of the remaining segments into a shorter string (see Table I, example C-3). In nearly all cases the remaining segments are recombined in a linear, left-to-right way, rather than being reordered within the linear string; the most typical pattern is that the onset of the first target syllable combines with the rhyme of the second target syllable. We have 30 telescoping errors in the corpus, and in all cases the tone spoken in the error utterance is the tone of one of the original syllables; no blended or hybrid tones occur. In 15 cases (50 %), the tones of the two target syllables are the same, suggesting that telescoping is more likely to take place between adjacent syllables with the same tone. The patterns involved in the other 15 cases, where the tones of the two target syllables differ, are somewhat equivocal, in ways similar to the blend errors discussed above.

Of these 15 cases, nine show the pattern where the spoken tone is that of the spoken rhyme, as shown in (29a) and C-3 in Table I. In the other six cases, the tone planned for the first syntagmatic slot is spoken, while the segments of the rhyme (three cases, e.g. (29b)) or the whole syllable (three cases, e.g. (29c)) of the second planned syllable are spoken (although in three of these latter cases, both rhymes have the same underlying vowel, as in (29c)).

(29) a. I: ʐɻ51-pən21 ly21-jow35 → E: ʐɻ51-pən21 **ljow35**
Japan trip (*onset* [l] + *rhyme* [jow35])
'a trip to Japan' → (ungrammatical)

b. I: kɑw21-pu51-xɑw21 ɕjɑŋ51-kɤ ta51 fej35 p^hwɑ35 →
maybe like-CL big fat woman
E: kɑw21-**pɑw51** ɕjɑŋ51-kɤ ta51 fej35 p^hwɑ35
'Maybe (you) will end up looking like a big fat woman.' →
(ungrammatical)

c. I: tɕjɛn21-fɑŋ55 pan51 an51 → E: tɕjɛn21 **pan55** ...
prosecutor make case
'The prosecutor dealt with this case.' → (ungrammatical)

One could argue from these errors that there is no evidence that tones are linked to segments in phonological representations, as the tone from the first or second planned syllable is nearly equally likely to be spoken in the error utterance. However, another possible interpretation would be that tones are associated with rhymes in phonological representations, but that telescoping errors can occur at two different points in the planning process, lexical insertion and phonological processing, which would give

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

rise to two different patterns of error. On this hypothesis one would argue that the first type of error (29a) occurs when an incorrect positional level frame is planned, so that there are not enough slots in the syntagmatic frame to support all the selected lexical items. In this case, when the lexical items are being inserted into the deficient frames, two syllables planned to be uttered in sequence will in fact become blended and inserted into the same syntagmatic slot. Because at this point in the derivation tones are not yet delinked from their rhymes, tones linked to deleted rhymes will also be deleted, and tones linked to inserted rhymes will be spoken in the erroneous utterance.

In the other case, exemplified by (29b, c), one could argue that the correct positional level frame was selected, the phonological representations of syllables were assigned to correct syntagmatic slots, and then the segments and tones delinked, in the usual manner. The error then takes place during the assignment of phonetic form; specifically, segments including a TBU are omitted. Because the tones are now associated with syntagmatic positions, the deleted TBU does not cause a tone to be omitted. Remaining segments are reassociated leftward to unfilled CV slots, and when tones become re-linked to segments for speech production, the leftmost tone is linked to the new segments in its associated CV slots. The rightmost tones as well as the unfilled CV slots are deleted, as they have no segmental realisation. If this were the case, it would explain why there are no errors involving rhymes from the syllable planned for the first syntagmatic slot, but tones from the second slot, since the rightmost tones would always be stranded and deleted.

Clearly, there is not enough evidence from these errors to settle the issue of whether tones are unlinked or linked to segments in phonological representations. The most conservative conclusion is that there is no evidence that tones need to be linked in underlying representations. However, to the extent that the telescoping and blend data support a theory of an underlying TBU, they implicate the rhyme in this role.

4.2.3 *Errors of tone.* Turning now to errors involving tones themselves, we find more evidence for the independence of tones and segments during the assignment of phonetic form. In every case where tones themselves are substituted (N = 78), exchanged (N = 4) or added (N = 1), the segments are not affected. Consider the following two tone-substitution errors:

(30) a. I: t^{h}wej55-tɕjɛn**51** xan35 → E: t^{h}wej55-tɕjɛn**51** xan**51**
recommend letter
'letter of recommendation' → (ungrammatical)

b. I: jow21 i**35**-ts^{h}i**51** tɕi55-tu55 tɕjɑw51 →
have one time Christian religion
E: jow21 i**35**-tsi**51** tɕi**35**-tu**51** tɕjɑw51
'There was once a time when Christianity...' → (ungrammatical)

In (30a), tone 4 [51] is perseverated and substituted for tone 2 [35], and the segments are not affected. In (30b), the tone pattern of the first

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

compound [i35-ts^{h}i51] 'one time' is perseverated and substituted for the tone pattern of the second compound [tɕi55-tu55] 'Christian', without affecting the segments.

In the one tone-addition error, a tone is added to an unstressed grammatical particle with underlying neutral tone, and the segments are not affected.

(31) I: mej21 san55-kɤ ʂi**35** fən55-tʂoŋ55 →
every three-CL ten minutes
E: mej21 san55-kɤ**35** ʂi**35** fən55-tʂoŋ55
'every 30 minutes' → (ungrammatical)

The four errors which we have classified as tone exchanges are actually somewhat ambiguous in that they all involve a tone 2 [35] exchanging with a tone 3 [21], as in (32).

(32) I: tɕi51 ɕjɑw**21**-lan**35** → E: tɕi51 ɕjɑw**35**-lan**21**
Ji Xiao-Lan
'person's name (Xiao-Lan Ji)' → (ungrammatical)

In each case, an alternative analysis would be that the tone 3 is either anticipated or perseverated, and substituted for a tone 2, giving a sequence of two tone 3s. This will cause the first tone 3 to become a tone 2 by tone sandhi, giving the identical surface result.

Looking specifically at the issue of tone sandhi, we find that there is ample evidence for the claim that tone sandhi is a completely productive phonological process which occurs after errors take place, since the sandhi rule is never violated in surface forms created by errors. First, if the 'reversals' discussed above are in fact substitutions, then they can be taken as evidence that tone sandhi takes place after tones have become associated with the segments with which they will occur in the spoken utterance, i.e. after the errors have occurred. Other tone-substitution errors, such as that shown in (33), confirm this analysis. In this error the tone 3 [21] from either [jow] or [tʂi] is substituted for the tone 4 [51] of [maj], causing the sequence [jow21 maj21]; tone sandhi causes the first tone 3 to become tone 2. We found no exceptions to this pattern.

(33) I: na**35** jow**21** maj51 pɑw51-tʂi21 →
where sell newspaper
E: na**35** jow**35** maj**21** pɑw51-tʂi21
'Where is the place selling newspapers?' → (ungrammatical)

Finally, clear evidence for the productivity of tone sandhi comes from lexical substitution and omission errors, in which the process operates on the resultant erroneous string, as illustrated in (34).

(34) a. I: jɑw51 tɕ^{h}y51 tʂɑw**21** ɕyɛ35-ɕyɑw51 →
want to go look for school
E: jɑw51 tɕ^{h}y51 tʂɑw**35 lɑw21-ʂi55**
want to go look for teacher
'(I) want to look for a school.' → '(I) want to look for a teacher.'

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

b. I: wɔ**35** tsɑw21 xwej35-laj35 → E: wɔ**21** ∅ xwej35-laj35
I early back come
'I came back early.' → 'I came back.'

In example (34a), the word [lɑw21-ʂi55] 'teacher' is substituted for [ɕɥɛ35-ɕjɑw51] 'school', causing a sequence of two tone 3's, so that the first tone 3 changes to tone 2 [35]. In example (33b), the word [tsɑw21] is omitted, allowing the word [wɔ21] to surface with its underlying tone. We found no exceptions to the claim that tone sandhi applies to the syntagmatic string after all errors have occurred.

4.2.4 *The role of tone in syntagmatic planning.* In order to address the question of whether or not tone is an organising factor in syntagmatic planning, similar to stress in English (see §2.3.2 above), we identified every contextual phonological error involving segments where either a unique source could be identified, or if there were more than one source, where all sources had the same tone (N = 399). We then counted the number of errors in which the target syllable and source syllable had each of the four tones. The results of this count are shown in Table II. It can

Source \ Target	Tone 1 (55)	Tone 2 (35)	Tone 3 (21)	Tone 4 (51)	Neutral tone	*total*
Tone 1 (55)	**25**	15	15	27	0	82
Tone 2 (35)	20	**25**	16	17	0	78
Tone 3 (21)	22	20	**16**	28	0	86
Tone 4 (51)	42	17	36	**49**	0	144
Neutral tone	4	0	2	3	**0**	9
total	113	77	85	124	0	399

[Table II. The number of phonological errors in which the target and source segments come from syllables with each tone. Numbers in boldface are cases in which the segments come from syllables with identical tones.]

be seen from this table that the syllables containing the target and source segments involve the same tone in only 29 % of the errors (115 out of 399 errors). While this is slightly greater than the 25 % expected by chance (if each tone occurs with approximately equal frequency), it is far fewer cases than the 95 % same-stress syllables found by Shattuck-Hufnagel (1983: 125) for English. Furthermore, it is clear from Table II that the largest number of errors in our corpus involve segments from syllables with tone 4. In order to see whether this was an accident of our data, we performed an informal count of tone frequency in Mandarin texts, looking at a passage from a newspaper, a novel, and a comic book. We found that tone 4 is in fact more frequent than the other three tones (tone 4, 32 % (N = 71); tone 1, 20·5 % (N = 46), tone 2, 19·5 % (N = 44), tone 3, 19 % (N = 43), neutral tone, 9 % (N = 20)). This elevated frequency explains why we

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

found a somewhat larger percentage of same-tone syllable errors than expected by chance, since 42·5 % of the same-tone syllables involved tone 4, which is the most frequent tone. The outcome of this analysis confirms that tone is not an organising factor in syntagmatic phonological planning, but is, as predicted, a property of lexical organisation.

4.3 Are contour tones in Mandarin unitary or sequences?

4.3.1 *Tones and segments.* Our data strongly support the hypothesis that Mandarin contour tones are unitary underlyingly, and we find no evidence from the data to support a claim that at some level of representation they consist of a sequence of level tones.

One of the most striking results of this study is the similarity between tones and segments in speech errors. First, both tones and segments can be moved, added, substituted and omitted as single units, and act as single units in the majority of errors. Recall that no lexical blend or telescoping errors produced hybrid tones. In the one tone-addition error ((31) above), a tone 2 [35] was added to the neutral tone syllable as a whole unit, and in the error where the tone pattern from one compound was copied onto another compound ((30b) above), the entire tone pattern was substituted as a unit. Second, we noted above that Shattuck-Hufnagel & Klatt (1979) found that consonants in English are involved in segmental errors in numbers which are proportional to the frequency of occurrence of the consonant in the language. We looked at the 68 errors in our corpus involving contextual substitutions of tones, and found a similar pattern, as shown in Table III. Because tone 4 is the most common tone in Mandarin,

Source \ Target	Tone 1 (55)	Tone 2 (35)	Tone 3 (21)	Tone 4 (51)	Neutral tone	*total*
Tone 1 (55)	*n/a*	3	7	4	0	14
Tone 2 (35)	4	*n/a*	6	9	1	20
Tone 3 (21)	0	4	*n/a*	8	0	12
Tone 4 (51)	8	7	7	*n/a*	0	22
total	12	14	20	21	1	68

[Table III. The number of cases in which each tone was involved in contextual tone-substitution errors, either as target or as source tone.]

we predicted that tone 4 would be involved in tonal errors more often than the other three tones, and in fact we find that 31·5 % of the errors involve tone 4 as either the source or target tone compared to tone 1: 19 %, tone 2: 25 %, tone 3: 23·5 % and neutral tone: 1 %. Third, Fromkin (1973) finds that speech errors are most likely to involve the production of real segments, and we find, similarly, that speech errors in Mandarin involve real tones. We found exactly one tone error that involved a non-Mandarin

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

tone, and no errors in the entire corpus in which a non-Mandarin segment was produced. Thus we have a great deal of evidence to support the hypothesis that tones and segments are treated as the same kinds of entities by the speech production planning system.

If we consider tones and segments to be similar kinds of units in the processing system, then there are at least two ways we could assess the internal structure of contour tones, in comparison to segments. The first would follow from the use by Yip, Duanmu and others of the term 'feature' for the pitch-related units within contour tones; the obvious comparison, then would be between these tone 'features' and segmental features such as [voice]. The second type of comparison arises from the fact that the posited pitch units within contour tones are in fact sequential rather than simultaneous (as segmental features are); thus a better comparison might be between contour tones and complex segments such as affricates and diphthongs, which are comprised of a sequence of phonetically different units mapped onto the same root node (Anderson 1978, Yip 1989). We will discuss each of these two possibilities in detail.

4.3.2 *Tone features.* Since we have shown that tones are treated as the same kinds of units as segments in the processing of Mandarin, we are justified in using the same criteria, laid out in §3.2.2 above, for determining whether a tone or segment error is contextual or non-contextual. In our corpus there are 353 phonological errors involving single consonants or vowels other than telescoping errors. Of these there are 90, or 25·5 %, in which there is no unit in the utterance identical to the error unit; these are the potential non-contextual errors. Of these there are 44 errors (12·5 %) in which a featural analysis is possible, including 10 involving the feature [aspiration] (see example C-1 in Table I), nine involving place of articulation (as in example (16a) above) and 25 which involve interactions between vowels and glides. In these latter errors, the source and target or error differ on the feature [syllabic], but originate from the same underlying vowel. (We take these errors as strong support for the hypothesis that glides derive from vowels.) For example:

(35) a. I: xwe**j**35-tɕ^{h}**y**51-lɤ → E: xwe**j**35-tɕ^{h} **i**51-lɤ
go-back-PERF
'went back' → (ungrammatical)

b. I: tʂ^{h}ɑŋ21-t**i**51 → E: tɕ^{h}**j**ɑŋ21-t**i**51
place
'place' → (ungrammatical)

Our analysis of (35a) is that the target vowel [y] was erroneously produced as [i] under the influence of the glide [j] which comes underlyingly from the vowel /i/; it could also be an error involving the feature [round] with [y], becoming unrounded in the context of [ej]. Our analysis of (35b) is that the vowel /i/ was anticipated and added after the onset consonant in [tʂ^{h}ɑŋ]; however, since this caused a sequence of two vowels to occur in the nucleus, the first became a glide, which caused the retroflex affricate to palatalise, producing the error form [tɕ^{h}jɑŋ]. In the other 46 (13 %) non-

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

contextual errors, an analysis involving features is not supportable by our criteria.

Looking now at the tone errors, we find that there are 83 tone errors in our corpus, with a total of 10 (12 %) errors in which the tone spoken in the error does not otherwise occur in the utterance (i.e. there is no source tone). Analysing tone errors in terms of tonal features is somewhat more problematic than looking for segmental feature errors. As discussed above, a segmental error was considered featural only if there was a potential source segment in the utterance which differed by only two features from the target segment, and was no farther than two syllables distant from the error. This is appropriate for segments because there are a relatively large number of features which could potentially differentiate any two segments. However, for the tones of Mandarin, if we consider the pitch heights which make up the four tones as being potential tone features, then there are only four possible tone features: levels 1, 2, 3 and 5; similarly, following Yip's analysis, there is only one register feature [upper]. Because there are so few tone features, we can predict that all of the tone features will regularly occur in any random utterance, and thus suggesting that a non-contextual tone was produced under the influence of one of these tone features somewhere else in the utterance would be tautologous. Therefore in order to constrain our analysis, we looked at the non-contextual tone errors for evidence that these tone levels or the register feature had spread from immediately contiguous tone features of contiguous syllables; this could involve tone features from both contiguous syllables or only one contiguous syllable. Note that this analysis is in line with autosegmental accounts which do not allow spreading across other tone features.

Before turning to our tone-error data, we need to discuss the fact that since errors nearly always result in a possible unit in a language, not all sequences of tones would allow spreading such that a real Mandarin tone would result. For example, while the sequence [35 xx 55] (where xx stands for any tone) would result in the allowable tone [35 55 55] if tone features were spread from both contiguous syllables, the sequence [35 xx 21] would result in the disallowed sequence [35 *52 21]. Similarly, a sequence [35 35] would result in the allowed sequence [35 55] if just the second tone feature of the first tone spread rightward to the first slot in the second tone, but [21 55] would result in [21 *15]. Finally the sequences [# xx 51] or [35 xx #] (where # is an utterance boundary) would result in the allowed tone [# 55 51] or [35 55 #] if the contiguous tone feature were spread to both tone features of xx, but the sequence [# xx 35] or [21 xx #] would result in disallowed sequences of [# *33 35] or [21 *11 #]. We calculated all possible combinations of tone sequences, and found that legal tones would be produced by spreading in 25 % of bidirectional medial spreadings, in 25 % of rightward unidirectional spreadings, in 12 % of leftward unidirectional spreadings and in 50 % of edge spreadings, for an overall total of 21 % possible spreading environments.

Evidence for the register feature [upper] would be found if errors occurred in which the target [hl] sequence remained unchanged, but the

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

feature [upper] shifted its feature value. The only candidate for this type of error would be a shift of tone 3 [−upper, hl] to tone 4 [+upper, hl] in the context of a tone 1 [+upper, hh] or tone 2 [+upper, lh] (but not tone 4, since this would be a contextual error). A tone 4 could shift to tone 3 by switching the register feature, but since tone 3 is the only inherently [−upper] tone in Mandarin, there would always have to be a tone 3 in the utterance, and thus this would be counted as a contextual substitution of tone 3. Any other tones which could be produced by a change in the register feature are not Mandarin tones (e.g. tone 1 [+upper, hh] 55→ [−upper, hh] *22). Thus we would look for non-contextual errors involving a shift of tone 3 to tone 4 in the environment of tone 1 or 2.

Returning to our non-contextual tone errors, we found that of the 10 errors, only two could be explained in terms of spreading. The first looks like a classic case of tone spreading:

(36) I: ɚ51 ʂɨ35 tɕjow21 xɑw51 → E: ɚ51 ʂɨ35 tɕjow**55** xɑw51
two-ten-nine number
'number 29' → (ungrammatical)

In this case one could argue that the error tone [55] came from a spreading of the second half of the preceding tone [35] and the first half of the following tone [51]. The second case produced the only non-Mandarin tone in the corpus.

(37) I: noŋ51 i51 ɕjɑw21 tʂʰa35-tʂʰi35 →
use one little teaspoon
E: noŋ51 i51 ɕjɑw**213** tʂʰa35-tʂʰi35
'using one little teaspoon' → (ungrammatical)

This appears to be an anticipatory addition of just the level [3] portion of the following tone [35], producing the non-Mandarin [213]. There were no non-contextual errors which could be explained by a spreading of the [upper] feature from contiguous syllables.

However, none of the other non-contextual tone errors was amenable to a tone-featural analysis. For example:

(38) I: tɑw21-jow35 xaj35 ʂwɔ55 → E: tɑw21-jow**51** xaj35 ʂwɔ55
travel guide also say
'the travel guide also said' → (ungrammatical)

In this error there is no non-vacuous source for the error tone [51]. In four errors, half the error tone is the same as an adjacent tone, but the other half is different from both the adjacent tone and the target tone, as in example (39):

(39) I: ni21-mən tɕjow51 tʂɨ21 xwej51 →
you-PL just only can
E: ni21-mən tɕjow51 tʂɨ**55** xwej51
'you only can' → (ungrammatical)

In this error, one could argue that the second half of the error tone [55] was caused by a spreading from the following tone [51], but there is no

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

way to explain why the first half of the tone [55] would be level [5], since it is adjacent to a level [1] pitch, and was originally a level [2] pitch.[6] The more supportable analysis of such errors is that they are simply a non-contextual substitution of one whole tone for another. Finally, in one error a tone changes to become less like the contiguous tonemes:

(40) I: ɚ51-ʂi35-ljow51 ʂow55-taw51 →
two-ten-six receive
E: ɚ51-ʂi35-ljow**21** ʂow55-tjɑw51
'receive (it) on the 26th' → (ungrammatical)

In this case a spreading would have produced [55], but in fact the dissimilar tone [21] is substituted for [51].

If tones were indeed being treated as similar to segments by the speech processing mechanism, we would expect approximately the same percentage of feature errors for tones as we found for segments. For segments the figure was 12·5 %, but if we accept the two errors (36) and (37) as our only clear evidence for tone features, this amounts to only 2·5 %, a much lower figure. We would argue that the tone substitution in (36) is simply a non-contextual error, and the fact that it can be explained as a spreading is merely a coincidence, since it is the only error out of the set of 10 non-contextual errors which allows this explanation. Similarly we would argue that the error in (37) is a low-level phonetic anticipation of the onset pitch of the following tone; the fact that only one error of this type occurred in our corpus suggests that it is an anomaly and should not carry much theoretical weight (indeed this error was made during very rapid colloquial speech). Finally, if we consider all 10 tone errors as non-contextual, then the percentage of pure non-contextual errors for segments and for tones is nearly identical (13 % *vs.* 12 %).

One possible flaw in the above reasoning is our assumption, based on similarity of error behaviours, that tones are being treated like segments by the processing mechanism. We used this assumption to limit our search for tone feature errors to the non-contextual tone substitutions. What would happen if we abandon this assumption and look at all 83 tone errors for evidence of tone features? In order to pursue this issue we analysed all 83 tone errors in terms of the following two factors: (i) Could the tonal environment allow a spreading (of either one or two features) which would result in a legal Mandarin tone (either a medial [55] or [51], or an edge [55])? (ii) Could the error tone in fact be the result of the spreading of one or two tone features?

We found that of the 83 tone errors, 10 provided an environment in which both contiguous tone features could have spread to cause the error,

[6] One could argue that the entire tone 3 [21] was deleted, and the contiguous tone feature [5] spread from the right into both empty tone feature slots. But then one would have to come up with a principled reason why in this case a single tone feature spread into two preceding empty slots, whereas in example (34) each empty tone slot was filled by the immediately contiguous tone feature, one from the right and one from the left.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

and of these, there were seven cases where the resultant tone could have been caused by spreading, as exemplified in (41):

(41) I: tʰa**55**-mən tʂi**21** jɑw**51** → E: tʰa**55**-mən tʂi**55** jɑw**51**
he-PL only need
'they only need ...' → (ungrammatical)

In this case, rather than arguing that the preceding tone [55] has been perseverated and substituted for the tone [21], one could argue that the second feature of the preceding tone [55] and the first feature of the following tone [51] have spread to cause the error tone [55] to substitute for the target tone [21] (with the preceding tone feature spreading through the neutral tone syllable [mən]). Furthermore, we found that there were 24 errors in which the environment provided the possibility of a single tone feature spreading and causing an allowable Mandarin tone, and of these there were 10 cases where the error tone could be the result of spreading, as in (42).

(42) a. I: tʰa**55** jɛ**21** ʂi**51** → E: tʰa**55** jɛ**51** ʂi**51**
he also is
'he is also' → (ungrammatical)
b. I: ʂwaj**55**-tɑw**21** → E: ʂwaj**55**-tɑw**55**
fall down
'fall down' → (ungrammatical)

In example (42a), rather than analysing this as tone [51] being anticipated and substituted for tone [21], one could take this as a case where the second feature of the first tone [55] has spread and substituted for the first feature of the second tone [21], causing the error tone [51]. In (42b) one could argue that this error does not involve the perseveration/substitution of tone [55] for tone [21], but rather the perseveration of just the second half of the tone [55], which spreads to both tone slots in the syllable at the edge of the phrase. In the remaining 49 cases, however, there was no environment which could have allowed the error to be caused by tone spreading.

What do these numbers tell us about tone features? Looking at them another way, we see that there were 34 errors (41 %) in which the environment for feature spreading was present. Our calculations discussed in §4.3 above showed that in fact only 21 % of tonal environments in Mandarin allow for legal tones being created as the result of spreading, so there was a much higher possibility in our error utterances for spreading to occur. However, of these 34 errors, only 17 (20·5 % of 83) could actually be analysed as tone spreadings, whereas the other 18 cannot be the result of tone spreadings under any analysis. This means that of the 83 tone errors, 66 (79·5 %) cannot be analysed as tone spreadings. Furthermore, there were 18 cases in our data where the error tone was a medial tone [55] or [51] or an edge tone [55] (the tones which can be created by tone spreading) which were not the result of spreading. Of the 35 errors in which the error tone was one which could be created by spreading, in 30

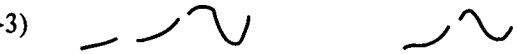
https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

cases there was another tone [55] or [51] in the utterance, and of the five non-contextual instances, only one can be analysed as a tone spreading ((38) above). This strongly suggest that analysing tone errors as whole-tone substitutions, additions or exchanges accounts for much more of the data than a featural analysis, and thus is a far more explanatory solution.

The reader may have noticed that the percentage of tone errors which can be analysed as tone-spreading errors (20·5%) is higher than the percentage of segmental errors analysed as being caused by featural influences (12%). However, recall that segmental errors were only analysed as being featural if there was no possible whole-segment source in the utterance. If we were to reanalyse all of the segmental errors, both contextual and non-contextual, in terms of whether the error *feature* occurred in a nearby segment, this analysis would allow us to analyse 87% of the segmental errors as featural, since by definition in all contextual errors there is a nearby segment with the source feature. As discussed above, we have good reason to believe that segmental errors typically involve whole segments even when featural analyses are possible. In order to argue that the featural analysis of tones is preferable to the whole-tone analysis even when a whole-tone source is available in the utterance, we would have to essentially ignore the many similarities between the behaviour of tones and segments in speech errors, and demonstrate that tones have an entirely different relationship to their features. But then this would leave us with the difficult task of explaining why the frequency of whole tones, not tone features, correlates with the frequency of individual tones in errors. Furthermore, we would have the contingent burden of explaining why in nearly 80% of the errors in our corpus, no featural explanation is possible under any analysis. For these reasons we claim that the whole-tone analysis is strongly supported by our data.

Finally, further support for the unitary status of Mandarin tones comes from a more detailed look at the telescoping errors. Odden (1995: 446) points out that in many African languages, when segments are deleted by phrase-level rules, the tones associated with deleted TBUs are not deleted; rather they are incorporated into the tone patterns of the remaining segments. Although telescoping is not strictly comparable, as it is an error rather than a phrase-level rule, one might hypothesise that in Mandarin telescopings, both tones of the originally planned syllables are kept during planning, but since there is now only one TBU available for realising the tones, a phonetically simplified (specifically shortened) tone results, which has as its starting point the starting pitch of the first tone, and as its ending point the ending pitch of the second. This is in fact the case for English telescoping, where the intonation pattern planned for the phrase is spoken in a condensed form over the telescoped phrase, as shown in (43) (Jaeger 1992).

(43)



https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

In Mandarin telescoping errors in which the two target syllables have the same underlying tone, this hypothesis cannot be tested since both the deletion and the shortening hypotheses would lead to the same results. However, we looked at the 15 telescoping errors in which the two target syllables had different underlying tones, and found that in seven cases, the shortening hypothesis could explain the results, as shown in (44a), where the spoken tone [35] could be a shortening of the sequence [35 55] rather than a realisation of the tone [35] planned for the final syntagmatic slot. But in the other eight cases such an analysis was not possible. In six cases (e.g. (44b)) a shortening would have resulted in a legal tone, but in fact the spoken tone simply reflects one of the two planned underlying tones; in the other two cases a shortening would have resulted in an illegal tone (e.g. (44c)). There was no correlation between a possible shortening analysis and whether the telescoping was of the 'rhyme' type or the 'positional' type.

(44) a. I: **wej35-an55** tɕjow21-tɕjow35 → E: **wan35** tɕjow21-tɕjow35
Wei-An uncle
'Uncle Wei-An' → (ungrammatical)

b. I: kən55 **tɕjɛ35 i55**-mwɔ35-i35-jɑŋ51 →
with sister one-size-one-kind
E: kən55 **tɕi55** mwɔ35-i35-jɑŋ51
'exactly the same as my sister' → (ungrammatical)

c. I: ʂɤ51-**tʰ**wan**35** t**ɑw**51 → E: ʂɤ51-**tʰɑw35** [*31]
club arrive
'arrive at the club' → (ungrammatical)

Although the shortening hypothesis can account for some of the data, there were no examples of telescoping which required the shortening hypothesis, for example, a sequence of [51 35] resulting in [55], where neither of the target syllables had tone [55] underlyingly. On the other hand, the deletion of whole tone hypothesis accounts for all of the data, including all examples which the shortening hypothesis cannot account for. These findings from telescopings again point out differences between African languages and Mandarin, differences which are easily accounted for if African contour tones are in fact underlyingly sequences, but Mandarin contour tones are underlyingly unitary. They also reinforce the claim that tones are lexical rather than being part of the syntagmatic organisation of the utterance, as they behave differently from intonation patterns in English, which do truncate when segments are lost in telescoping.

4.3.3 *Complex segments.* The second possibility for probing the internal structure of contour tones is to argue that they are in fact structurally more similar to complex segments such as affricates and diphthongs than they are to simple segments, and thus an analysis of errors involving affricates and diphthongs will be more revealing. Affricates are typically analysed as being a sequence of two consonantal feature bundles which map onto a

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

single consonantal root node, and diphthongs are a sequence of two vocalic units which map onto a vowel root node, so these segments could clearly be seen as analogous to Yip's analysis of contour tones illustrated in (11b) above. There are several points which need to be made about such a comparison before we turn to these errors.

First, Mandarin has unambiguous affricates, which function as single units in the consonantal system (see Table I). However, Mandarin does not have diphthongs. As discussed in §2.1, the pre- and postnuclear glides are best analysed as sister elements to vowels under the nucleus node, but vowel + glide sequences do not function as single units in the phonological system. Thus, the fact that glides act independently of vowels in Mandarin errors, as we will illustrate briefly below, has no actual bearing on the tone argument. Secondly, from a phonetic point of view, there are some clear differences between contour tones and affricates, in that affricates are always articulated with two distinct gestures, which produce a discontinuous acoustic signal (stop + burst, followed by a fricative), whereas contour tones are produced with a laryngeal gesture which causes a continuous acoustic signal. In fact, contour tones have more in common phonetically with true diphthongs than with affricates, in that diphthongs are produced with a gradually shifting gesture, so that it is difficult to see on a spectrogram where the 'steady-state' portion of the vowel ends and the glide begins (in the case of offglides). Third, there is an important phonological difference between affricates and diphthongs compared to contour tones, which has to do with constraints on the sequence of elements. Affricates are always, by definition, a stop followed by a fricative (just as prenasalised stops are always a nasal stop followed by an oral stop, and so on). While diphthongs can involve onglides or offglides cross-linguistically, in most languages only one or the other is allowed (for example, in Korean only onglides are allowed, whereas in English all but one diphthong involves offglides). This is in fact similar to the tone systems of some African languages: in Margi (Nigeria) [hh], [ll] and [lh] tonal sequences are allowed, but *[hl] is disallowed (Pulleyblank 1997). However, in Mandarin, as in most Asian tone languages, there are no constraints on the sequence of pitch elements: Mandarin and other Chinese languages have rising, falling and level tones, Thai has rising, falling, dipping and level tones, and so on. It is clear that contour tones in Asian languages do not show the same sequencing constraints as affricates (and other complex consonants), tones in some African languages and, to some extent, diphthongs. In sum, there are clear phonetic and phonological differences between contour tones and affricates; while contour tones and diphthongs have more in common, Mandarin does not have true diphthongs, which makes the comparison difficult. Nevertheless, a look at errors involving affricates and vowel + glide sequences in Mandarin may be instructive. In particular, if it is found that affricates never split into component parts in Mandarin errors, this can be argued to show that the lack of clear tone-spreading or splitting errors discussed above may be

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

irrelevant, since the inability to split in an error may simply be a property of complex phonological units, segments and tones alike.

Errors involving glides confirm that glides clearly do not form diphthongs with vowels; furthermore, there are no ordering constraints on vowel + glide sequences, again indicating that these are not diphthongs. There are 41 errors in the corpus in which glides are involved in errors separately from adjacent vowels, and 22 in which vowel + glide acts as a unit. For example:

(45) a. I: t^{h}o**w**35-pən55 → E: t^{h}o**w**35-p**w**ən55
project-run
'flee to freedom' → (ungrammatical)

b. I: t^{h}a**j**35 ta51 wa**j**51 wən35 → E: t^{h}a**j**35 ta51 wa**j**51 we**j**35
Taiwan University foreign literature
'foreign literature department in Taiwan University' → (ungrammatical)

c. I: t^{h}o**ŋ**35-ɕyɛ35 two55-pan51 → E: t^{h}o**ŋ**35-ɕyɛ35 to**ŋ**55-pan51
classmates most
'most of my classmates' → (ungrammatical)

d. I: me**j**35 jow21 k^{h}ɤ51 → E: me**j**35 jow21 k^{h}e**j**51
not have class
'does not have a class' → (ungrammatical)

In the first two examples, the glide acts as an independent unit: in (45a), the postnuclear glide [w] is perseverated and added to the prenuclear glide slot in the following word. In (45b), the postnuclear glide [j] is perseverated and substituted for the nasal coda of the following word; the vowel then shifts to its correct allophone [e]. In the second two examples, the glide functions as a unit with the vowel: in (45c) the rhyme [oŋ] is perseverated and substituted for the rhyme [wɔ]; in (45d) the rhyme [ej] substitutes for the rhyme [ɤ]. If we were to use errors involving vowel + glide as a comparison with tonal errors, they would provide very strong support for the claim that Mandarin contour tones are unitary, since the behaviour of tones in errors is very different from that of vowel + glide sequences, which divide freely. However, as indicated above, because these are not true diphthongs, this comparison is not valid. It would be more valid to compare Mandarin tonal errors to errors involving diphthongs in a language such as English, although this suffers the methodological problem of comparing behaviour of speakers of different languages. Errors which break up diphthongs in English are in fact very rare, although the second author has recorded errors such as [hórm.wɚk] for [hó^{w}m.wɚk] *homework*, where the rhotacised vowel from the second syllable is anticipated and its non-syllabic form substituted for the offglide of the diphthong. However, a more systematic comparison of contour tones and diphthongs awaits further study, and will be particularly convincing with data from speakers of a language which has both.

Turning to affricates, we note that it is commonly reported that affricates do not split up in speech errors (Stemberger 1983). However, in

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

order to investigate this question in Mandarin, we looked at the 28 errors which we classified as non-contextual onset-consonant substitutions, looking to see whether some of these 'non-contextual' errors could in fact be analysed as affricate splits (recall that affricates only occur in onsets). In fact, eight of these errors could be analysed as involving individual segments from affricates; in five cases one element of the affricate has been omitted, as illustrated in (46a), and in three cases the source segment could be one element of a contextual affricate, as illustrated in (46b).

(46) a. I: tʰa55 tɕjow51 xwej51 ʈ͡ʂan51-tʂɤ təŋ21 →
he then will stand-PROG wait
E: tʰa55 tɕjow51 xwej51 ɕan51-tʂɤ təŋ21
'Then he will keep waiting.' → (ungrammatical)
b. I: tswɔ35-tʰjɛn55 lɑw21-ʂɨ55 tɕ͡ʝɑw55-tɤ →
yesterday teacher teach
E: tswɔ35-ɕjɛn55 lɑw21-ʂɨ55 tɕ͡ʝɑw55-tɤ
'Yesterday what the teacher taught.' → (ungrammatical)

In (46a), the stop portion of the affricate [tʂ] could have been omitted in dissimilation to the other [t]'s in the utterance (a 'masking' error); in (46b), the error segment [ɕ] could be an anticipation from the affricate [tɕ]. (Similar errors have in fact been recorded in English by the second author: 'Gèt twò páckages of [frò^{w}.ʒən] ... fròzen spínach', where the [ʒ] in *packages* [pʰæ.kɨ.ʤɨz] could be the source of the error.) Thus, contrary to previous reports, complex segments can show some decompositional behaviour in errors. If we accept these as errors showing the decomposition of affricates into two distinct elements, then we can again use this as evidence against there being internal structure in contour tones. In our data there were 49 errors involving the substitution, addition or omission of affricates, and eight (16%) show possible decompositional behaviour, whereas of the 83 tone errors, only two (2·5%) can be analysed as decomposition errors in a straightforward manner.

In sum, evidence from both the 'featural' analysis and the 'complex segment' analysis converges on the finding that contour tones in Mandarin show no evidence of having internal structure, i.e. of being composed of a sequence of level tones.

5 Conclusions

In this paper we have presented an analysis of 788 slips of the tongue made by native speakers of Mandarin, in terms of the light they shed on various issues regarding the representation and processing of tone. All but the first of the questions we posed at the beginning of the paper have received clear answers:

(a) While tones are clearly represented along with segments in UR, our data do not provide unambiguous evidence on the question of whether they are unlinked or preassociated in UR; there is some suggestion that

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

they may be associated with rhymes, but all the data are compatible with their being unassociated. When one lexical item is substituted for another, added in a new location or misordered in a syntagmatic string, its tone remains intact; when a lexical item is deleted, its tone is deleted. When two words are blended into a single phonological form during lexical insertion, one of the two underlying tones is spoken in the error, usually that of the spoken rhyme, although this evidence is not overwhelming. The fact that two words involved in a lexical substitution or blend error have the same tones twice as often as expected by chance suggests that tones are an important component of the phonological organisation of the lexicon.

(b) After lexical insertion, during the mapping from phonological representations to phonetic form, tones either remain unlinked, or become delinked from their segments; segments and tones are involved in errors independently from each other during this process. After tones and segments have been assigned their final ordering (i.e. after all phonological errors have taken place), tone sandhi is applied to the resultant string.[7] In telescoping, when TBUs are deleted, two patterns emerge: in some cases the tone of the retained rhyme was also retained, suggesting that these telescopings occurred during lexical insertion, and are akin to the blend errors in which tones remain linked to rhymes. In other cases the tone of the first syntagmatic slot was retained, regardless of which rhyme was deleted, suggesting that these deletions occurred during phonological processing.

(c) Tone is not as critical an organising factor for syntagmatic planning as stress is in languages such as English, in that in only 32 % of segmental errors, the target and source syllables contain the same tones, which is only slightly more than expected by chance.

(d) Our evidence strongly supports the hypothesis that Mandarin tones are unitary and are not made up of sequences of tones or tone features at any level of representation or processing. First, there were no tone splittings or hybrid tones produced by word-blend or telescoping errors; while there were some telescoping errors that could be analysed as involving tone spreadings, none required this analysis. Secondly, in the only tone-addition error in the corpus, a whole tone was added to a neutral tone syllable; similarly, in the one error where tones from a source compound were substituted for the tones of a target compound, the entire tone pattern of the source was substituted intact. Third, there were only two cases of non-contextual tone errors which allowed a tone feature-spreading analysis, and none which allowed a register feature-spreading analysis. Taking all 83 tone errors together, contextual and non-contextual, only 17 (20·5 %) could be explained by a tone-spreading analysis, while 73 (88 %) could be explained as simple whole-tone contextual errors, and nine of the other 10 (11 %) as whole-tone non-contextual errors. The fact that 44 of the non-contextual segmental errors (12·5 % of

[7] Other phonological processes, such as palatalisation, gliding, allophone of vowel, etc., also take place after new orderings are assigned.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

the total number of segmental errors) are best analysed as feature errors (i.e. [place], [aspiration], [syllabic]) supports the hypothesis that segments are made up of phonetic features. But the fact that only two of the non-contextual tonal errors (2·5% of the total number of tonal errors) can be analysed as tone-spreading errors supports the hypothesis that tones in Mandarin are fundamentally unitary. Finally, comparing the contour tones to the analogous complex segments, affricates, we found that the stop + fricative units split apart in 16% of the errors involving affricates, suggesting that the absence of splitting errors found with tones is not simply a function of the unitary nature of complex phonological units which project from the same root node. Thus we claim that native speakers of Mandarin represent and process each tone phonemically as a distinctive but unitary combination of pitch and contour.

The answers to our first two questions provide evidence that the general autosegmental theory of tones in terms of phonological representation and assignment of phonetic form discussed in the introduction to this paper has clear psycholinguistic support. In this case the processing facts and the theoretical hypotheses map simply onto each other, demonstrating that the phonological theory has a good possibility of having cognitive status. On the other hand, the answer to our fourth question suggests that phonological theories which require that all contour tones in every language must be represented as a sequence of level tones underlyingly may be missing an insight into the possible underlying differences among tone languages. We conclude by suggesting that if phonological theories of tone intend to reflect cognitive facts, their theoretical constructs must be flexible enough to encompass both types of underlying tone system.

REFERENCES

Anderson, S. (1978). Tone features. In Fromkin (1978). 133–175.

Bao, Z. (1990a). *Fanqie* languages and reduplication. *LI* **21**. 317–350.

Bao, Z. (1990b). *On the nature of tone*. PhD dissertation, MIT.

Beckman, M. (1986). *Stress and non-stress accent*. Dordrecht: Foris.

Berg, T. (1987). *A cross-linguistic comparison of slips of the tongue*. Indiana University Linguistics Club.

Bock, K. & W. Levelt (1994). Language production: grammatical encoding. In M. Gernsbacher (ed.) *Handbook of psycholinguistics*. San Diego: Academic Press. 945–984.

Bogers, K., H. van der Hulst & M. Mous (eds.) (1986). *The phonological representation of suprasegmentals*. Dordrecht: Foris.

Caplan, D. (1992). *Language: structure, processing, and disorders*. Cambridge, Mass.: MIT Press.

Carr, P. (1993). *Phonology*. London: Macmillan.

Chao, Y-R. (1930). A system of tone-letters. *Le maître phonétique* (3rd series) **45**. 24–27.

Chen, C-Y. (1984). Neutral tone in Mandarin: phonotactic description and the issue of the norm. *Journal of Chinese Linguistics* **12**. 299–333.

Chen, J-Y. (1993). A corpus of speech errors in Mandarin Chinese and their classification. *The World of Chinese Language* **69**. 26–41.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

Cheng, C-C. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.
Chomsky, N. & M. Halle (1968). *The sound pattern of English*. New York: Harper & Row.
Chung, R-F. (1989). *Aspects of Kejia phonology*. PhD dissertation, University of Illinois at Urbana-Champaign.
Clements, G. N. & J. Goldsmith (eds.) (1984). *Autosegmental studies in Bantu tone*. Dordrecht: Foris.
Cutler, A. (ed.) (1982a). *Slips of the tongue and language production*. The Hague: Mouton.
Cutler, A. (1982b). The reliability of speech error data. In Cutler (1982a). 7–28.
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.
Dell, G. & P. Reich (1981). Stages in sentence production: an analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior* **20**. 611–629.
Demuth, K. (1992). The acquisition of Sesotho. In Slobin (1992). 557–638.
Derwing, B. (1973). *Transformational grammar as a theory of language acquisition*. Cambridge: Cambridge University Press.
Duanmu, S. (1990). *A formal study of syllable, tone, stress and domain in Chinese languages*. PhD dissertation, MIT.
Duanmu, S. (1994). Against contour tone units. *LI* **25**. 555–608.
Erbaugh, M. (1992). The acquisition of Mandarin. In Slobin (1992). 373–455.
Fromkin, V. (1973). The non-anomalous nature of anomalous utterances. In V. Fromkin (ed.) *Speech errors as linguistic evidence*. The Hague: Mouton. 215–242. Appeared originally (1971) in *Lg* **47**. 27–52.
Fromkin, V. (ed.) (1978). *Tone: a linguistic survey*. New York: Academic Press.
Fromkin, V. (ed.) (1980). *Errors in linguistic performance: slips of the tongue, ear, pen, and hand*. New York: Academic Press.
Gandour, J. (1977). Counterfeit tones in the speech of Southern Thai bidialectals. *Lingua* **41**. 125–143.
Garrett, M. (1975). The analysis of sentence production. In G. Bower (ed.) *Psychology of learning and motivation*. Vol. 9. New York: Academic Press. 137–177.
Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (ed.) *Language production*. Vol. 1: *Speech and talk*. London: Academic Press. 177–220.
Garrett, M. (1984). The organization of processing structure for langauge production. In D. Caplan, A. Lecourse & A. Smith (eds.) *Biological perspectives on language*. Cambridge, Mass.: MIT Press. 172–193.
Garrett, M. (1988). Processes in language production. In F. Newmeyer (ed.) *Linguistics: the Cambridge survey*. Vol. 3: *Language: psychological and biological aspects*. Cambridge: Cambridge University Press. 69–96.
Goldsmith, J. (1976). *Autosegmental phonology*. PhD dissertation, MIT. Published 1979, New York: Garland.
Goldsmith, J. (1979). The aims of autosegmental phonology. In D. Dinnsen (ed.) *Current approaches to phonological theory*. Bloomington: Indiana University Press. 202–222.
Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.
Goldsmith, J. (1995a). Phonological theory. In Goldsmith (1995b). 1–23.
Goldsmith, J. (ed.) (1995b). *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell.
Halle, M. & G. N. Clements (1983). *Problem book in phonology*. Cambridge, Mass.: MIT Press.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

Halle, M. & K. Stevens (1971). A note on laryngeal features. *MIT Quarterly Progress Report* **101**. 198–212.
Hombert, J.-M. (1986). Word games: some implications for analysis of tone and other phonological constructs. In Ohala & Jaeger (1986b). 175–186.
Hyman, L. (1973). Discussion paper on Sino-Tibetan tone. Paper presented at the 6th International Conference on Sino-Tibetan Languages and Linguistics, University of California, San Diego.
Hyman, L. (1986). The representation of multiple tone heights. In Bogers *et al.* (1986). 109–152.
Hyman, Larry M. (1993). Register tones and tonal geometry. In H. van der Hulst & K. Snider (eds.) *The phonology of tone: the representation of tonal register*. Berlin: Mouton de Gruyter. 75–108.
Jaeger, J. (1992). 'Not by the chair of my hinney hin hin': some general properties of slips of the tongue in young children. *Journal of Child Language* **17**. 335–366.
Jaeger, J., A. Lockwood, D. Kemmerer, R. Van Valin, Jr., B. Murphy & H. Khalak (1996). A positron emission tomographic study of regular and irregular verb morphology in English. *Lg* **72**. 451–497.
Katamba, F. (1989). *An introduction to phonology*. London: Longman.
Kiraly, G. (1996). *Hungarian slips on the icy slope of language*. Honors thesis, Department of Linguistics, SUNY, Buffalo.
Laubstein, A. (1987). Syllable structure: the speech error evidence. *Canadian Journal of Linguistics* **32**. 339–363.
Leben, W. (1973). *Suprasegmental phonology*. PhD dissertation, MIT. Published by Garland Press, New York.
Leben, W. (1978). The representation of tone. In Fromkin (1978). 177–219.
Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge, Mass.: MIT Press.
Lin, Y-H. (1989). *Autosegmental treatment of segmental process in Chinese phonology*. PhD dissertation. University of Texas at Austin.
Lin, Y-H. (1990). Prenuclear glides in Chinese. Paper presented at the Mid-America Linguistics Conference, University of Kansas.
Mohanan, K. P. (1995). The organization of the grammar. In Goldsmith (1995b). 24–69.
Moster, D. (1991). Slips of the tongue and pen in Chinese. *Sino-Platonic Papers* **22**. 1–45.
Mowrey, R. & I. MacKay (1990). Phonological primitives: electromyographic speech error evidence. *JASA* **88**. 1299–1322.
Newman, P. (1986). Contour tones as phonemic primes in Grebo. In Bogers *et al.* (1986). 175–193.
Odden, D. (1995). Tone: African languages. In Goldsmith (1995b). 444–475.
Ohala, J. (1986). Consumer's guide to evidence in phonology. *Phonology Yearbook* **3**. 3–26.
Ohala, J. (1995). Experimental phonology. In Goldsmith (1995b). 713–722.
Ohala, J. & J. Jaeger (1986). Introduction. In J. Ohala & J. Jaeger (eds.) *Experimental phonology*. Orlando, Fl.: Academic Press. 1–12.
Pike, K. (1948). *Tone languages*. Ann Arbor: University of Michigan Press.
Pulleyblank, D. (1986). *Tone in Lexical Phonology*. Dordrecht: Reidel.
Pulleyblank, D. (1997). Optimality Theory and features. In D. Archangeli & D. Langendoen (eds.) *Optimality Theory: an overview*. Oxford: Blackwell. 59–101.
Sapir, E. (1921). *Language: an introduction to the study of speech*. New York: Harcourt, Brace & Co.
Shattuck, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. MacNeilage (ed.) *The production of speech*. New York: Springer. 109–136.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press

Shattuck-Hufnagel, S. & 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*. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale-Ecole des Hautes Études en Sciences Sociales. 139–162.

Slobin, D. I. (ed.) (1992). *The crosslinguistic study of language acquisition*. Vol. 3. Hillsdale, NJ: Erlbaum.

Stemberger, J. (1983). *Speech errors and theoretical phonology: a review*. Indiana University Linguistics Club.

Wan, I-P. (1996). *Tone errors in Mandarin Chinese*. MA thesis, SUNY, Buffalo.

Wan, I-P. (1997). The status of prenuclear glides in Mandarin Chinese: evidence from speech errors. *CLS* **33**. 417–428.

Wan, I-P. (forthcoming). *Mandarin phonology: evidence from speech errors*. PhD dissertation, SUNY, Buffalo.

Wan, I-P. & J. Jaeger (1996). Tone errors in Mandarin Chinese. Paper presented at the 5th International Conference on Chinese Linguistics, National Tsin Hua University.

Wan, I-P. & J. Jaeger (1997). The representation of tone in Mandarin: evidence from speech errors. Paper presented at the 71st Annual Meeting of the Linguistic Society of America, Chicago.

Wang, W. (1967). Phonological features of tone. *IJAL* **33**. 93–105.

Wang, Z. (1996). Features as phonological primitives: evidence from Beijing Mandarin. Paper presented at the 70th Annual Meeting of the Linguistic Society of America, San Diego.

Williams, E. (1976). Underlying tone in Margi and Igbo. *LI* **7**. 463–484.

Woo, N. (1969). *Prosody and phonology*. PhD dissertation, MIT.

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.

Yip, M. (1980). *The tonal phonology of Chinese*. PhD dissertation, MIT.

Yip, M. (1989). Contour tones. *Phonology* **6**. 149–174.

Yip, M. (1995). Tone in East Asian languages. In Goldsmith (1995b). 476–494.

https://doi.org/10.1017/S0952675799003668 Published online by Cambridge University Press