PROBLEMS OF HIERARCHY AND INDETERMINACY IN MANDARIN PHONOLOGY

SAMUEL E. MARTIN

Yale University

0. **Introduction.** It seems appropriate to greet Y. R. Chao with a paper on two aspects of phonologic analysis which have attracted increasing attention in the years since he first directed interest toward them in his paper The non-uniqueness of phonemic solutions of phonetic systems (1934). I shall refer to these two aspects as problems of hierarchy and problems of indeterminacy, and I shall discuss them in relation to the structure of Mandarin Chinese. This paper has developed from an earlier unupublished one (titled Subsidiary phonemic systems in Mandarin) written while I was studying under Chao at the University of California. The paper has since benefited from suggestions by F. K. Li, K. L. Pike, N. C. Bodman, K'un Chang, Morris Swadesh, and André Martinet.

1. **Structure and patterns.** We all agree, I think, that a language HAS a phonologic structure; we are not agreed on the inclusiveness and the internal organization of this structure, nor on the determinability of its make-up, i. e. on the "uniqueness" of our various analyses. In addition, we are often confused by what we mean by "a language". From the wide usage of this term, we can arrive at four different types of "Mandarin phonemic systems":

(1) The pattern of distinctions in the idiolect of any given native speaker who enjoys a high degree of mutual intelligibility with all other speakers of what we may call Northern Mandarin. (This could presumably be extended to include Southwest and Southern Mandarin by saying "a moderate degree of intelligibility", and it could even be extended, if we wished, to the kind of "local-tradition" Mandarin fluently spoken by many Chinese from non-Mandarin areas.) We do not necessarily expect the phonologic system of any one such idiolect to agree in its entirety with any other.

(2) The MINIMUM pattern of the distinctions common to all the idiolects (called by Hockett and others COMMON CORE). This might prove to be identical with the system of the idiolect making the fewest distinctions; or, it might be found to contain fewer distinctions than any single idiolect.

(3) The MAXIMUM pattern found in all the idiolects (called by Hockett and others OVERALL PATTERN). This might prove to be identical with the system of the idiolect making the most distinctions; or, it might turn out to contain more distinctions than any single idiolect.

(4) The pattern of the idiolect of any given speaker ADJUSTED to either the minimum or the maximum pattern, by the subtraction or addition of the extra distinctions. That is, an individual system plus (or minus) a fragmentary sub-system, thought of as together forming a whole, but one with an internal organization possibly different from that of the minimum (or maximum) pattern itself.

2. **Subsidiary systems.** Of a somewhat different nature is the hypothesis offered by Fries and Pike in "Coexistent Phonemic Systems", Language 25.29–50 (1949), who raise the possibility that a single idiolect may contain within itself one phonemic system in complete form, and another system of fragmentary nature to account for aberrant items drastically limited in distribution. Since I assume that each idiolect is a distinct entity, I find it necessary to modify their suggestion in the following way: The only phonemic system in "complete" form is an all-inclusive one sufficiently expanded and reorganized to cover the additional data for which the fragmentary system was postulated. But internal organization of this system may reveal (or conceal) subsidiary systems of varying stability and importance, probably brought into being by the shifting sands of linguistic change.

The analyst at work often faces the dilemma "Should I conceal some of the data to reveal pattern, or should I conceal pattern in order to include all of the data?" It seems to be axiomatic that the more data acquired (and the more closely it is examined), the more ragged becomes the pattern. The inescapable conclusion: language patterns are more complex than linguists' pictures of them might suggest. Perhaps an analogy with cartography is valid: no map includes all the facts about a given piece of terrain, but only those facts thought relevant for the purpose of the map.

The problem of handling certain distinctions in Mandarin is relevant here:

(1) Is the so-called "raised 3d tone" (here, with Hockett, called "the 5th tone"—not to be confused with the "entering tone" of Southern Mandarin) a separate phoneme from tones 2 and 3?

(2) Are there two mid-vowel distinctions (i. e. two kinds of "e"), and how are these to be handled?

(3) Are there two low-vowel distinctions (i. e. two kinds of "a") and how are these to be treated?

These three problems occur only as the result of certain morpheme combinations; the functional load of the distinctions (when made) seems to be quite low, and it is questionable whether they occur in the speech of a majority of of Northern Mandarin idiolects. Related to the mid-vowel problem is the question:

(4) Can [ü] be treated as the sequence /iu/ or must it be regarded as a separate phoneme?

3. **Identification.** The determinability of ULTIMATE PHONOLOGIC CONSTITUENTS (the distinctive "components, features, aspects" or "qualities" which bundle together to make up the "phonemes")—both what the constituents are and when they are present—seems to hang on our notions of what has been called "phonetic similarity" and can be re-interpreted as "auditory identification (with articulatory feedback)". Although phonetics is usually discussed in purely articulatory terms, we know that the phonologic entities are primarily auditory units; similarity must be thought of in perceptual terms. But our knowledge of the auditory-perceptual nature of phonemes is largely limited to articulatory observation and acoustic measurement, so that we are forced to talk in articulatory or acoustic terms. When we speak, however, of one phonetic segment as being more "similar" to a certain segment than to another, we run the danger of basing our conclusions on observations which may not always correlate with the auditory realities; segmentation itself is under considerable suspicion (see Language 32.702).

Questions of identification show up in Mandarin in several places:

(5) Do the series of "palatal" initials (roughly represented by j_i, ch_i, sh_i, y_i) and the series of "retroflex" initials (roughly j_r, ch_r, sh_r, r_r) share revelant components with each other? Do either (or both) share relevant components with the "alveolar" series (d, t, s, dz, ts) or with the "velar" series (g, k, h, and "zero")?

(6) Are the phonetic phenomena called the "zero" (vowel-onset) initial to be regarded as conditioned or significant? Are the onsets of initial i-, u-, and r- phonetically parallel and phonologically identical?

(7) Are the aspirated stops [p'], [t'], etc., to be treated as a sequence of the unaspirated [b̥], [d̥], etc., plus a non-velar allophone of /h/? The initial /h/ is velar; could the velar quality be attributed to the "zero" initial, i.e. can [xɑ] be analyzed /ɣhɑ/?

In order to present all these problems in their setting, I offer below a comparison between a composite of older treatments and a new analysis of the phonemic system of a kind of "maximum Pekinese Mandarin". In the new analysis, an alveolar vowel z is postulated as a relevant phonetic feature; the affrication involved in certain phoneme sequences (j_i, ch_i, j_r, ch_r, dz, ts) is treated as conditioned. Hockett does just the opposite with these two features of Mandarin articulation. The present treatment may be no improvement; it remains to be proved, however, that it has any less validity.

4. **Mandarin phonemics.** Since considerable phonetic description of the phonemes can be found elsewhere,[1] allophones are not discussed except when especially relevant.[2] By convention, syllables are separated by space, with the stress written before (in Roman numerals), and the tone after (in Arabic numerals). The basic point of departure is to consider tone as

1. E.g.: Y.R. Chao, Mandarin Primer (Cambridge 1948): Id., Introduction on Pronunciation, Mathews' Chinese-English Dictionary, Revised American Edition, pp ix-xvii (Cambridge 1944); Id. and L.S. Yang, Concise Dictionary of Spoken Chinese (Cambridge 1947); L.M. Hartman, The Segmental Phonemes of the Peiping Dialect, Language 20. 28-42 (1944); C.F. Hockett, Peiping phonology, Journal of the American Oriental Society 67. 253-67 (1947); Bernhard Karlgren, A Mandarin Phonetic Reader in the Pekinese Dialect, with an introductory essay on the pronunciation (Stockholm 1918). Further bibliography can be found in these works; the literature is extensive. An acquaintance with the Hartman and Hockett papers will help in following the arguments in this paper. See also Hockett's Peiping Morphonemics, Language 26. 63-85 (1950). The Hartman paper and both Hockett papers are reprinted in Readings in Linguistics, edited by Martin Joos for the ACLS (1957); Chao's article on Non-uniqueness will also be found there. Please note that I use "ng" in this paper as a digraph for the single phoneme /ŋ/. The phonetic symbols, y, w, and ü are used in the Americanist tradition.

2. I should like to point out some examples of rapid speech supporting the biphonemic status of the retroflex and palatal consonants: ian0 reng0 for hian0 hreng0 'Mr.'; IIpu2 r0 IIIpu2 r0 for IIpu2 hr0 IIpu2 hr0 'No no'; IIpɑu4 rɑng0 for IIpɑu4 hrɑng0 'in the newspaper'; pu0 r0 tɑu0 and pu0 rɑu0 for IIIpu4 IIItrl IItɑu0 'don't know' used parenthetically. Another point which has been made, but deserves greater emphasis, is the extent to which the tonal range decreases as the stress diminishes, so that at each step of stress reduction the contours grow closer until contour distinction is obliterated altogether into the neutral tone with no stress. In a similar way, vowel quality distinctions narrow in range, approaching or coalescing with the unstressed mid vowel.

the nuclear unit. That is, the "syllable" is defined as any sequence of phonemes carrying a tone; the neutral tone is here treated as a phoneme, not just the absence of some other tone. Intonation and emphatic stress are peripheral to the main line of discussion, but it seems desirable to include a brief summary of the best available descriptions of these phenomena along with the other data. Much remains to be discovered about stress, juncture, and intonation.

It would be somewhat misleading to use the symbols j (or y) and w, together with the term "semivowel" for what is here written i and u. I wish to state this, because it reverses an earlier attitude of mine. Phonetically, these phonemes nearly always receive a distinctly vowel-like articulation; the front vowel has two important vocoid allophones: [ɪ] after the alveolar and labial consonants (and also after a low vowel), [i] elsewhere. There is no need to set up a specific nuclear "vowel" for each syllable if we regard the tone as the nuclear unit. The actual distribution of acoustic energy and of auditory prominence over the elements of the syllable (shall we say "tonable"?) varies with tone, stress, and segmental environment. In any event, the difference in prominence is considerably less marked than the terms "semivowel, vowel" would suggest to a native speaker of English.[3]

5. **Intonation phonemes.** Following the data of Chao,[4] there are two "successive" phonemes of intonation; these are added on the tone of the last syllable of a phrase, and perhaps correspond to the "terminal junctures" in the terms of Trager and Smith. In addition there are four "simultaneous" phonemes which affect the tonal range throughout a phrase or phrase-group with increasing effect toward the end; these probably correspond to Trager and Smith's "voice qualifiers." The effects of the "successive" phonemes on the allophones of the various tones are given below, following Chao, on an ascending numerical scale with 1 as the lower limit of the voice, and 5 the normal—i. e. colorless—upper limit.

3. For a more detailed discussion cf. Karlgren, op. cit. 58-61 "Distribution of stress within the syllable". See also Martin, The Phonemes of Ancient Chinese 22 (Suppl. 16, JAOS 1953).

4. Y. R. Chao, Tone and Intonation in Chinese, Bulletin of the Institute of History and Philology 4.121-34 (1933); Id., A Preliminary Study of English Intonation (with American Variants) and Its Chinese Equivalents, The Ts'ai Yuan P'ei Anniversary Volume, Supplementary Volume I of the Bulletin of the Institute of History and Philology (1932). A summary of some of this material may be found more accessible in K. L. Pike, Tone Languages (Ann Arbor 1948). especially 85-6.

/X/	Rise.	/1/	[55:]⟶[56:]
		/2/	[35:]⟶[36:]
		/3/	[214:]⟶[216:]
		/4/	[51:]⟶[513:]
		/0/	conditioned⟶[5:]
/Y/	Fall.	/1/	[55:]⟶[551:]
		/2/	[35:]⟶[351:]
		/3/	[214:]⟶[2141:]
		/4/	[51:]⟶[5121:]
		/0/	conditioned⟶[1:]

The raised-3d tone (tone /5/) never occurs final in a phrase.

The simultaneous phonemes of intonation are:

/A/ General raised level of pitch.

/B/ General lowered level of pitch.

/C/ Widening of range.

/D/ Narrowing of range.

The first two are usually (but not always) accompanied by the last. These simultaneous phonemes of intonation could be written before a phrase; the successive ones could be written after the last syllable. Other special interjection intonations exist, but have been excluded from this presentation.

6. **Stress phonemes.** There are three (but possibly only two) stress phonemes; lack of stress is here not considered a phoneme itself, but the absence of other stress phonemes, though this point may very well deserve reconsideration (since "unstressing" functions as a morpheme).

/I/ Emphatic. (The phonemic form of the morpheme of special emphasis.)

/II/ Normal.

/III/ Reduced. A still weaker allophone occurs when preceded by another reduced stress and followed by a normal stress.

The phonemic status of Stress III is perhaps open to question, at least for many speakers who do not seem to make the subtle emphasis-distinctions that Hockett has reported. If morphological considerations are taken into account (as Pike might be willing to do), the two allophones represented by Stress III would seem to be in complementary distribution with Stress II. The contrast, when present, seems to be between verb-object and adjective-noun constructions: IItɑ3 IIIren2 'hit people', IIIhɑu3 IIren2 'good people'. We

might prefer to speak of a different kind of syllable juncture in the two cases.

7. **Tone phonemes.** There are six tone phonemes, and the principal allophones can be described somewhat as follows (using the system that divides the pitch range from 1 to 5):

/1/ [55:] (with an elegant minority of speakers, [54:] before /0/)
/2/ [35:]
/3/ [214:] before pause and [21:] elsewhere
/4/ [53:] before /4/ and [51:] elsewhere
/5/ [24:] (see below)
/0/ usually short, with the pitch determined by the environment[5], but [3:] in rare occurrences after pause.

The neutral tone is treated as a phoneme /0/. Since the pitch is almost entirely determined by environment, it can be argued that the occurrence of the neutral tone is just the absence of any tone phoneme. Note, however, that tone /0/ is not always accompanied by lack of stress; interjections have accompanying stress, and are to be considered in the neutral tone with some intonation phoneme usually in addition.

Tone /5/ appears only as an alternant of a morpheme whose basic tone is /3/. When such a morpheme precedes another morpheme whose basic tone is /3/, the first morpheme changes to tone /2/ (unless itself preceded by tones /1/ or /2/; in this case, the morpheme changes to tone /1/). But if the preceding tone-3 morpheme is accompanied by the emphatic stress /I/, then the tone change is to /5/. Thus IImai2 IIImɑ3 'buy a horse' is homonymous with IImai2 IIImɑ3 'bury a horse' at ordinary stress; but with emphatic stress Imai5 IIImɑ3 'BUY a horse' contrasts with Imai2 IIImɑ3 'BURY a

5. For a complete description, cf. Mandarin Primer 90-1, 108-9. In summary:
After Tones 1 and 4 ...[1]
After Tones 2 and 5....[3]....[2]....[1]
After Tones 3 ....[4]....[3]....[2]....[1]
But with the following exceptions:
After Tones 1 and 2....[5]....before Tone 4.
After Tones 3 ....[2]....before Tones 1, 2, 4.
One addition might be made to the remarks in the last paragraph of page 27 in Mandarin Primer: Synonymous binomial compounds—that is, words consisting of two syllables each of which have essentially similar meaning—usually have the last syllable in the neutral tone: daw.luh 'road', perng.yeou 'friend', kann.jiann 'to see', shii.huan 'to like' (and huan.shii 'be glad'), kuay.leh 'happy', shyng.wei 'actions'. Cf. Koosaka Zyun-iti: (Sinpen) Sina-go hatuon ziten [(New) Chinese Pronunciation Dictionary], Tokyo 1942, p. 217. Denzel Carr in "A Characterization of the Chinese National Language", Bulletin de la Société Polonaise de Linguistique 3.3 (1931), p. 64.9, points out that there are numerous exceptions.

horse'. Tone /5/ also appears before syllables in the neutral tone /0/. Here it contrasts with tones /2/ and /3/ as in these examples: IImai2 tiɑr0 'bury some', IImai5 tiɑr0 'buy some', IImai3 te0 'bought'. (The basic tone forms are mai2 'bury', mai3 'buy', tiɑr3 'some', te0 [particle].)

Except for the emphatic cases before all tone–3 syllables, the only reason that we need to recognize a fifth tone is that Tone /3/ tends to be raised, yet kept distinct from Tone /2/, before certain neutral-tone syllables but not before all of them. With the exception of a limited list of reduplicated relationship terms such as IIkie3 kie0 'elder sister', it can be said that Tone /3/ changes to Tone /5/ before a neutral-tone syllable which is an allomorph of a morpheme basically Tone /3/. The tone has been lost with the loss of stress, but there remains a sort of morphophonemic "memory" of its presence which raises the preceding Tone /3/. We see this in the examples IIhiɑu5 kie0 'young lady', IItɑ5 ni0 'hit you'; cf. IIt*ɑ*3 ne0 'hit, you-see'. If it were not for these conflicting instances with neutral-tone syllables, the occurrence of Tone /5/ under emphatic stress before Tone /3/ could be regarded as in complementary distribution with the other allophones of Tone /3/.

8.0. **Vowel phonemes.** The vowel distinctions suggested by the present analysis can be charted as follows.

	Front	Central	Back
High	i	z r	u
Mid		e, e:	
Low	a		ɑ

The high vowel r is retroflex or retracted; the high vowel z is usually described as "apical". The auditory effect of this vowel is a sort of buzz, and it seems to be a phonetic type peculiar to China, though the /z/ of the Miyako dialect of the Ryukyus is quite similar. In Peking this sound occurs only after sibilant and affricate (making possible the treatment of each as a sequence of two phonemes as below); but in other dialects, this sound type is found also after other alveolars, e.g. [tz nz] 'geography' in the dialect of Chihsi in Anhwei. For the Peking dialect, the vowel is better called "alveolar", since the tongue-tip is actually touching the bottom teeth throughout the pronunciation of [sz], [dz] and [$t^{s}z$].

8.1. **Mid vowels.** The phoneme /e:/ represents a long mid vowel. It is usually a back diphthong [$ɤ^{\Lambda}$]; in the sequence /ue:/ we find [$u^{ɔ}ə$], in the

sequence /e:u/ we find [ᵊɔu];. The sequence /ie:u/ sounds like [iᵊɔu]; in other instances, when preceded or followed by /i/ or /iu/ the allophone is a slightly lowered and centralized [e:].

The phoneme /e/ represents a short higher-mid vowel [ə]. For those Americans who distinguish English /ɨ/ ("barred-I") from /ə/ ("shwa"), the Chinese phoneme probably sounds more like the high central vowel of their own speech, rather than the mid central one. The sound is less prominent [ᵊ] between high vowels and between /u/ and /n/. It is back [ʌ] before /ng/. Possibly the sequence here (as traditionally) treated as [ɨ̞ŋ]=/ing/ should be analyzed as [ɪᴬŋ]=/ieng/.

The distribution of the two mid-vowel phonemes can be shown as follows:

e:	ie:	ue:	iue:	e:m	—	—	—	—	e:i	ue:i
e	—	—	—	em	en	uen	eng	ueng	—	uei
e:u	ie:u	e:r	ie:r	ue:r	iue:r	e:ur	ie:ur	—	—	—
—	ieu	er	ier	uer	iuer	—	ieur	engr	iengr	uengr

Except for the underlined occurrences, /e:/ is in complementary distribution with /e/. Of the underlined occurrences, /ue:i/ and /ie:u/ regularly occur with tones 3, 4, and 5, and /uei/ and /ieu/ occur only with tones 1 and 2; but there are occurrences of /ue:i/ and /ie:u/ in tones 1 and 2 when a basically tone–3 morpheme changes to first or second tone in sandhi.[6] For the majority-type subdialect, there is no distinction between (i)(u)er and (i)(u)e:r in Tones 3, 4, 5, but there are contrasts (like kerl 'root' and ke:rl 'song') in Tones 1 and 2; some of these people retain /e:/ in tone–3 morphemes ending in /e:r/ when these morphemes undergo sandhi change to Tones 1 and 2, and others shorten the vowel to /e/. For a sizeable minority of speakers, these distinctions are not maintained in any tone (so that 'root' and 'song' are homonyms); for another minority they are said to exist for all tones.

All the relevant finals in -r can be analyzed as derivative compounds with the suffix which has the shape /r/ (after u, e, a) alternating with the shape /er/ (after morpheme-final r, z, and i not preceded by vowel); post-vocalic -i, -n drop automatically before adding the suffix, and -ng either drops (younger generation) or is pronounced simultaneously with the vowel and the following /r/.

6. E. g. γie:u3 'have'+kie:u3 'nine'+ping3 'cake' in (IIme:i2 γ ie:u0 IIIkie:u3 Iγuan4) IIIγ ie:u2 IIIkie:ul Iping3 '(not have 9-CHARACTER piece, but) have 9-CIRCLE piece (in Mahjong)'.

In addition there are the finals em and uem in contrast with e:m and ue:m. Aside from the words hrem2 'what', tzem3 'how', tzem4 'thus', nem3 'how', nem4 'so', nem0 'well then'—all of which in slower speech usually occur compounded with the neutral-tone linking particle e0 as in IItzem3 e0 'how'—the finals em and uem occur only as automatic morphophonemic alternants of en and uen before labials as in IIɣuem4 me0 'asked??', IIɣuem3 me0 'steady??' from ɣuen4 and ɣuen3. A contrast is provided by the compound ɣue:3+m=ɣue:m3 'we' in the phrase IIɣue:m3 me0 'us??' Note also IIhre:2 me0 'broken??' contrasting with IIhrem2 e0 'what?' The peculiar tonal behavior of this particular word, noted by Hockett, is exceptional regardless of the analysis. The tone is spread over both the stressed and unstressed syllables. This may be regarded as a unique distribution of Tone 2 (Hockett's treatment) or as a unique [34:] allophone of Tone 2 on the first syllable followed by the high allophone of the neutral tone on the second syllable; in final position this high allophone could be ascribed to the intonation phoneme /X/ (rise). But in non-final position this intonation could hardly be held responsible; however in non-final position hrem2 e0 does tend to be replaced by the simpler hrem2, e.g. in IIIhrem2 (e0) IItungl hi0 'what things'. As a matter of fact, the expected tonal allophones of hrem2 e0 do occur in certain utterances. Note the distinctions below:

1. IIIni3 IItr1 tau0 IIIɣue:3 IIIhrue:1 IIhrem2 e0 (X) [34:, 5:]
2. IIIni3 IItr1 tau0 IIIɣue:3 IIIhrue:1 IIhrem2 me0 (X) [35:, 5:]
3. IIIni3 IItr1 tau0 IIIɣue:3 IIIhrue:1 IIhrem2 me0 [35:, 3:]

The first sentence means 'Do you know what I SAID?' (Answer: 'You said such-and-so.') The last two both mean 'Do you KNOW what I said??' (Answer: 'Yes' or 'No.')

The more usual morphemic analysis of the word for 'what' would be hre2 plus a suffix me0, with hrem2 as an abbreviation containing a shorter form of the suffix. I find it somewhat simpler to think of hrem2 a single morpheme since hre2 does not occur alone; the ambisyllabicity of the /m/ is a product of the usual liaison of the interrogative particle e0. This treatment has the advantage of removing all stressed syllable-final occurrences of short /e/ except for the two cases noted below.

The finals e: and e are in complementary distribution for the most part; the long vowel occurs in stressed syllables, the short vowel in unstressed ones. Two occurrences of the short vowel in stressed syllables are IItre4

hr0 'this is' and IIne4 hr0 (alternating freely with IInɑ4 hr0) 'that is'. The form IItre:4 hr0 also occurs, but it has an additional sarcastic connotation. It may be that that there are speakers who say IItre:i4 hr0 phonetically parallel to IItre:i4 ke0 but I have not heard them.

[Note: Hockett's treatment is actually more subtle than my remarks indicate. He finds a number of quasi-dissyllabic "microsegments", and these could presumably include IItre4 hr0. But the phonetics of such microsegments are difficult to hear, since a "spread" of tones 3 and 4 over two syllables would ordinarily sound no different from the ordinary tones followed by a neutral-tone syllable; only the word for 'what' gives us the necessary pitch contrast. In other cases one has to rely on rhythm to reveal the lack of expected juncture.]

8.2. **Low vowels.** The phoneme /a/ is a front low vowel, somewhat higher and fronter in the final ian [iɛn], somewhat centralized [ɐ] before /i/ or /n/ in Tone 1. The phoneme */a/* is a non-front low vowel, about central when final, back elsewhere. The distribution of the two phonemes may be shown as follows:

ɑ	ia	ua	iua	ɑm	uɑm	—	—	—	—	ɑng	iɑng
—	—	—	—	am	uam	an	ian	uan	iuan	—	—
uɑng	iuɑng	—	—	—	ɑu	iɑu	ar	iar	uar	iuar	aur
—	—	ai	iai	uai	—	—	—	—	—	—	—
iaur	ɑngr	iɑngr	uɑngr								
—	—	—	—								

These phonemes present still fewer points of contrast than /e/ and /e:/. The low vowels would be in complementary distribution were it not for two things: (1) the assimilation of final -n after the low vowel to a following labial or velar; and (2) the reduced form of the suffix me0, or more strictly the two homonymous suffixes one of which means 'plural number of persons' and the other 'pause for speaker to think'. Compare for example: IIthaml pu0 IIthrl le0 [from thanl+pu0] 'He (polite) doesn't eat any more.' IIthɑml pu0 IIthrl le0 [from thɑl+m+pu0) 'They don't eat any more' or 'Him, he doesn't eat any more.' Or, since the polite 3d-person pronoun is in increasing disuse, a less minimal contrast: IItzam2 pu0...'We (inclusive) don't...'. Also IIthzam2 pu0 IIthrl le0 [from thzan2] 'The silkworms don't eat any more' versus IIthrɑm2 pu0 IIthrl le0 [from thrɑ2+m+pu0] 'Tea I don't drink any more.'

The velar assimilation is illustrated by IIhzangl ke0 'three' [from hzan1] in contrast with IIhrɑngl le0 'injured'. If the velar assimilation were the only difficulty, it would perhaps be possible to assign the final nasal in the first instance above to the phoneme /n/, since it is considerably more front than the usual allophone of /ng/ after /ɑ/. However the vowels have to be kept distinct because of the contrast with the labial assimilation, so the allophone of pre-velar [ŋ] is probably better described as /ng/ fronted by the vowel.

Some believe that this assimilation of final -n to a following labial or velar is complete only with the low vowel, and for moderately slow speech perhaps only when the following syllable is unstressed. But there is an assimilatory tendency with the mid vowel, too, and it is conceivable that the distinction between the two allophones [ʌ] and [ə] of /e/ may sometimes be more meaningful than that between the weakly articulated nasal finals of such a pair as IIItrengl IIkaul 'steamed sponge-cake' and IIItrenl IIkaul 'real tall'.

Hockett apparently heard thɑml 'they' and tzam2 'we (inclusive)' with the same vowel, but they are clearly distinct with speakers that I have heard. One solution to the assimilation problem is to regard as the sequence /an/ any low front vowel quality accompanied by nasalization (and disregard the labial aspect of the [m] segment as conditioned overlap with the following labial); any low back quality accompanied by nasalization can then be /ang/ or /am/ depending on the actual nasal segment itself. Aside from the lack of phonetic parallelism in this solution, it does not take care of the vowel distinction between thɑml 'they' and tzam2 'we (inclusive)' when followed by a non-labial, e. g. IIIthɑml IIɣie:u3 'they have' versus IIItzam2 IIɣie:u3 'we have'. One way out of this dilemma would be to posit a unique final cluster -nm in /tzanm/ 'we'; this has certain merits.

8. 3. **Two additional problems.** After preparing this analysis another problem came to my attention, when the informant produced this set of minimal pairs: kerl 'root', ke:rl 'song', kɑrl 'liver', kɑ:rl 'dried skin', kɑngrl 'small jar'. This set would seem to indicate a length distinction for the low vowel paralleling that of the mid vowel. The particular word kɑ:rl 'dried skin' may be peculiar to Tientsin (where the informant, Grace Wan, was born). Henry Fenn suggests an alternative example in the Peking word kɑ:rl meaning 'the "cat" used in playing "tip-cat" (IItɑ3 IIIkɑ:rl)'. The Peking

word for 'dried skin' seems to be IIkɑ1 pe(r)0, also pronounced IIkɑ2 pe(r)0, IIkɑ(:)r1 per0, and IIkɑ(:)r2 per0; the educated version is IIke:1 pe0. Robert Tharp suggests that this word may be etymologically related to an onomatope meaning 'crackling, crisp (as of celery or weather)'. Hockett's solution of the 'root'-'song' distinction is to regard vowel quality as the relevant feature, rather than length, and pin the difference on the vowel of the short form which he hears as "palatalized" (whatever that may mean—presumably higher and fronter than elsewhere). Hockett treats our kenl (also meaning 'root') as /ken/ but our kerl as /keir/. I doubt that many will follow him on this lack of parallelism in the phonetics. Incidentally, Hockett cites the word /kair[1]/ 'dried stuff' which is probably our word kɑ:rl 'dried skin' (one of his informants was also in Tientsin); if this is the same word, I heard the phonetic parallel to the 'root'-'song' distinction as just the opposite from what Hockett heard, perhaps because I was looking for length and he was looking for "palatalization".

Hockett (in Peiping Morphophonemics 64) calls attention to the fact that /ien/ and /ian/ are best treated as distinct in any kind of "overall pattern" for the Peking dialect. For some speakers the form [iɛn] is limited to the first two tones, and [ian] occurs in the others, so that there is complementary distribution, and there is undoubtedly a good deal of free variation within an idiolect (as there is for /er/ and /ɑr/); but many speakers seem to have individual items with one or the other sequence quite uncorrelated with tone, and I suspect that any Peking speaker could readily be trained to tell you which of the two he has just used even though he feels perfectly free to choose either one for many (perhaps all) items. This may be regarded as an "emergent" distinction, of very low functional load, and we can feel free to normalize all items to either /ien/ or /ian/ for etymological purposes—I believe that /ian/ has certain advantages.

There is no need to concern ourselves here with the problem whether /e:/ is a unit phoneme distinct from /e/, with or without a shared component, or whether it is two phonemes /e/+/:/ or /e/+/e/. This is largely a question of the criteria for handling length (cf. Language 32.691).

An alternative way to handle the difficulties reflected in the two phonemes /e/ and /e:/ was once suggested by F.K. Li. Hartman, in his analysis, posited a zero nuclear vowel with distinctive feature of syllabicity which we might write /-/, to serve as a tone carrier for his phonemes /s, c, ch/ when

these are accompanied by no other vowel (since Hartman's system does not include our /z/), and similarly for the other high vowels /i, u, r/ which Hartman regards as basically non-nuclear semivowels. Now, Li suggests extending this zero syllabic, the phonetic quality of which is determined quite strictly by environment, to cases of the short mid vowel (our /e/). The phonemic interpretation of the phonetic data in this way, amusingly enough, would be almost the reverse of the well-known interpretation of English syllabic r and n as /ər, ən/ or /ɨr, ɨn/. Certain difficulties are present in the solution. The position of the "zero" phoneme must be clearly indicated: cr- (= /tr/) must be distinguished from c-r (= /tzer/). (The difference could, of course, be indicated indirectly by position of the tonal markers: cr3 and c3r.) A complete omission of Hartman's "zero" (which he symbolizes as i, using j and w for our /i/ and /u/) does not lead to any ambiguity in Hartman's own system.

Note the structure of the following contrasts in the two treatments:

tz	c-
tr	cr-
tzer, tzen	c-r, c-n
tze:r	cer
γue:m	uem
γuem	u-m
γiu	iu
γieu	i-u
γie:u	ieu

The system with the minimal syllabic would be unable to accomodate the short mid vowel in IItre4 hr0, IIne4 hr0.

8.4. **The high front rounded vowel.** Hockett originally followed Hartman and others in treating [ü] as the sequence /iu/. Later he abandoned this in favor of a unit phoneme; there are contrasts between what he hears as something like [iou] and what he hears as something like [iu]. Some of these are produced by sandhi (as *yu jiing* 'oil well' versus *yeou jiing* 'there are wells' with the proper stresses); others may be of the "emergent" distinction sort that we found with /ian/ and /ien/. One apparently consistent example is that of two interjections that can be spoken with Tone 1: [iou] 'gee whiz!' and [diu] 'toot!' We are not faced with the same dilemma as Hockett, since we have two mid-vowel distinctions, and can treat the [iou] forms as /ie:u/

and the [iu] forms as /ieu/—and sure enough we start hearing the latter as [i^{ə}u]. So we treat [ü] as /iu/. The reason for the order /iu/ rather than */ui/ for this simultaneous sequence has to do with distribution (there is a [fu] but no*[fi] or *[fü]) and with allophones (/k/ is a palatal affricate before /i/ and /iu/ but not before /u/).

9.0 **Consonant phonemes.** Phonemic treatment of consonants in other analyses can be summarized somewhat as follows; this is a composite of Hartman, Hockett, and Chao.

	Labial	Alveolar	Alveolar Affricate	Velar
Unaspirated	p	t	c	k
Aspirated	ph	th	ch	kh
Nasal	m	n		-ng
Lateral		l		
Spirant	f	s		h
Voiced spirant				γ

9.1. **The palatal and retroflex series.** The palatal affricates and sibilant (j_i, ch_i, sh_i) have been treated as /ci, chi, si/. The retroflex affricates and sibilant (j_r, ch_r, sh_r) have been treated as /cr, chr, sr/. The palatal series is also in complementary distribution with the velar series (we find no *ki, *khi, *hi); the retroflex series is in complementary distribution with both the velar and alveolar series (we find no *kr, *khr, *hr and no *tr, *thr alongside /sr/). Previous analysts have chosen to class the palatal and retroflex consonants with the alveolar affricates and spirant because there seemed to be greater phonetic similarity in the manner of articulation. That is, affrication (or sibilance) is treated as the most relevant phonetic feature.

Since the present analysis includes the alveolar vowel /z/ it is possible to say that c and ch are rather /tz/ and /thz/; the palatal affricates are then in complementary distribution with only the velar series (we find conflicting distribution with /ti/ and /thi) and they must be /ki/ and /khi/. The retroflex affricates can be either /tr/ and /thr/ or /kr/ and /khr/; the better case is made for /tr, thr/ both on articulatory and auditory grounds. The significant phonetic feature of /t/ is apical articulation (or its auditory equivalent), with the tongue-tip retracted or retroflexed when followed by /r/, and with affrication when followed by /z/ or (to a lesser extent when

you listen closely) by /r/. The significant feature of /k/ is dorsal articulation, palatalized and affricated (slightly) before /i/.[7]

And the sounds s, sh_i, and sh_r (earlier treated as /s, si, sr/) will be analyzed as /hz/, /hi/. and /hr/ respectively. It is misleading to label the sound [s] as an apical sound in Mandrin; the tongue is very far forward, and the tongue-tip rests against the lower teeth. We can consider the sound to be dorsal (in our sense—"any part of the tongue except the tip", non-apical) with an allophone fronted to alveolar position by the following vowel /z/. We define /t, th, n, l/ as APICAL and only /z/ as ALVEOLAR; we define /k, kh, h, -ng, γ /as DORSAL in a broad sense, just as "labial" is used in a broad sense to include bilabial /p, ph, m/ and labiodental /f/. If we extend the definition of /h/ to include the aspiration of /ph, th, kh/, we may want to remove the phoneme from the dorsal column altogether. It can be argued that the tongue-tip is used in the same way in /hr/ as in /tr/, but at least there is no contact between the tip itself and the mouth; we can still say that the /h/ is dorsal in our broad sense of the word, and the retroflex articulation of the following /r/ simply overlaps.

For a clearer picture of what the suggested treatment means in practice, let us consider some typical syllables. The forms on the left are phonemic, those on the right phonetic.[8]

—	ta	—	ka	—	ɖaˈ	—	ǧaˈ
—	tha	—	kha	—	tʻaˈ	—	kʻaˈ
—	—	—	ha	—	—	—	xaˈ
tz	tza	—	—	tsʐ	tsaˈ	—	—
thz	thza	—	—	tsʻʐ	tsʻaˈ	—	—
—	—	hz	hza	—	—	sʐ	saˈ
ti	tiau	ki	kia	ɖɪ	ɖɪau	ɖʐi	ɖʐiaˈ

7. Assignment of the palatal allophones to the velar phonemes corresponds to the treatment of the velar "inserted" consonant in one of the three types of secret speech described in Chao's "Eight varieties of Secret speech", Bulletin of the Institute of History and Philolgy 2.312-54 (1931); cf. especially 321, paragraph 3. 1. This has long been the solution favored by Chinese linguists. The analysis of the retroflex affricate as /tr/ is an idea of my own, inspired by Chao's pronunciation drills.

8. The phonetic symbolization follows Chao's usage in the main. Use of the symbol for "palatal t" should not lead us to misconceptions; the sound in question could just as well be termed a "palatal k". The sibilant nature of the affrication is not so pronounced as a symbolization with varieties of s would indicate; this is true for palatal, retroflex, and alveolar alike, though in the latter case the vowel itself has a characteristic sibilance. Cf. remarks in Carr, op. cit. 54, where he attributes to Chao the suggested notation [k̂ç].

thi	thiɑu	khi	khiɑ	t'ɪ	t'ɪɑʊ	ʈɕ'i	ʈɕ'i̯ɑ˦
—	—	hi	hiɑ	—	—	ɕi	ɕi̯ɑ˦
tr	trɑ	—	—	ʈ̬ʂ̬ɪ	ʈ̬ʂɑ˦	—	—
thr	thrɑ	—	—	ʈ̬ʂ'ɪ	ʈ̬ʂ'ɑ˦	—	—
hr	hrɑ	—	—	ʂɪ	ʂɑ˦	—	—

9.2. **The voiced spirant.** The phoneme /ɣ/ has been suggested for the majority-type dialect of Peking before mid and low vowels.[9] The phonetic nature is that of a voiced velar fricative in free variation with a uvular squeeze and with zero (i.e. free vocalic onset), also often a glottal stop in stressed position. (A minority of speakers replace this phoneme with /ng-/ before open vowels.) The basis for setting up the phoneme as distinctive is the behavior of several vowel-initial particles in the neutral tone: ɑ0, e0, eu0. These particles tend to link with a preceding nasal or high vowel, so that the final phoneme of the preceding syllable has an ambi-syllabic quality, making possible contrasts such as IIthruan2 e0 'boat....' and IIran2 ɣer0 'but....'. Hockett's solution of this difficulty[10] is to assume gemination of a preceding nasal (thus: nɑ0, ngɑ0) and ignore the similar phonetic phenomena with the high vowels. An extension of this treatment to include the high vowels (where there is no contrast) would produce the following allomorphs of a typical particle: ɑ0, iɑ0, zɑ0, rɑ0, uɑ0, nɑ0, ngɑ0, mɑ0, Note that this introduces unique initial occurrences of /z/ and /ng/ with neutral-tone syllables only; such, however, is the treatment apparently favored by some Chinese grammarians who have used a unique combination of the National Phonetic Letters to write /zɑ/ in their discussions of Chinese grammar. (The allomorphs mentioned above can be said to be conditioned by the preceding phoneme; not so conditioned is the /iɑ0/ allomorph of this particular particle which follows the stressed mid and low vowels.) Chao's treatment, while

9. Cf. Chao. "The voiced velar fricative as an initial in Mandarin", Le Maître Phonétique 89 (January-June 1948). Chao limits the occurrences assumed to the more open vowels. There would seem to be a good case, however, for extending the application to /i, u/ as well, in view, of the fricative articulation of the high vowels when initial: ɣil [yi] 'one', ɣu3 [wu] 'five'. In initial position a more fricative variety of /r/ often occurs too. (The high vowel z does not occur initially.) There are no real minimal contrasts for the high vowels, but it would perhaps make a more consistent pattern to assume that /ɣ/ occurs before them as well. In this paper, I have followed this practice with respect to /i, u/ but not /r/.

10. Cf. C. F. Hockett and C. Y. Fang, Spoken Chinese 237-8 (New York 1944); C. F. Hockett, Peiping Morphophonemics 77-8. Hockett's treatment obscures the contrast between items like IIItz3 IIɣanl e0 'How about Tzyyan?' and IIItz3 IIɣanl ne0 'Oh you mean Tzyyan?!'; IIthɑm e0 'How about them?', IIthɑm me0 'Them??', and IIthɑ me0 'Him??'

perhaps more cumbersome in practice, would seem to do better justice to the phonetics.

9.3. **The aspirated initials.** It is possible to consider the aspirated stops and affricate as the unaspirated phonemes plus /h/. This treatment was rejected by Hartman largely for reasons of pattern; /h/ alone did not occur before /i/. But in the present analysis, the sequence /hi/ occurs. The other possible objections: (1) The aspiration of the stops is glottal, but /h/ is considered dorsal. (2) There are no other initial clusters. (3) The aspirated stops are fortis, the unaspirated lenis. (This is a peculiarity of Mandarin which sets it apart from languages like Cantonese and Fukienese.) There is no objection to analyzing the aspirated stop as the unaspirated phoneme plus aspiration, but there is some doubt whether the result is two phonemes or two components of one phoneme. I also raise the question whether initial /h/ should be considered a parallel case and treated as /γh/ (before mid and low vowels, but not before /i/ and /z/), with a broadening of the definition of /γ/.

10. **Syllable structure.** A recapitulation of the new phonemic analysis in the form of a syllable-structure chart is presented below. Any Chinese syllable can be plotted on the chart, reading from left to right, with the vertical lines separating sequence positions and the horizontal lines indicating general distribution classes. It is to be understood that only immediate distributional limitations are indicated in the diagram, and the intonation phonemes are omitted altogether.

Stress I, II, III
Tone 1, 2, 3, 4, 5, 0

f					
γ p, ph k, kh m n l	i	u	e e: a ɑ ?ɑ:		i m n
				u ng	r
h t, th	r z				

Each syllable contains as a minimum one tone and one of the phonemes with the double lines; sequences of these phonemes also occur within the syllable. In very informal speech, the phonemes /m, ng/ occur alone with tone: m3∼m0 'us'; ng0 'Oh no indeed'. (with various stresses and intonations). In this same range of marginal speech are the interjections [ɛ] 'hey!' and [ɔ] 'oh!' which introduce two additional vowels; these have been left out of the discussion. Inclusion of [ɛ] and [ɔ] would not necessarily wreak havoc with the system we have set up, since the allophones of the various finals which include the phonetic symbols [ɛ] and [ɔ] are actually somewhat centralized, while the interjections are not. Also ignored is the initial glottal fricative, voiced and voiceless in free variation in the exclamation [ɦɑˈ] or [hɑˈ], indicating exhilaration or satisfaction "whee!", and the alveolar click indicating admiration. Inclusion of all interjections would probably necessitate a glottal stop phonemically distinct from /ʔ/.

2.6. **Articulatory overlap.** Chao has observed that the Chinese phonemes show a kind of "simultaneity of compatible articulations". This is true not only of tones, as would be expected, but of the other phonemes as well. For example, the palatal position of /i/ is assumed with a preceding consonant and is pronounced quite simultaneously with a following /u/ (resulting in [ü]), the influence continuing (with or without the intervening /u/) to a following /e:/ which is fronted from its usual back value. The labial position for /u/ is held throughout a syllable like /trungl/ 'middle'. The vowel /z/ occurs simultaneously with a preceding consonant (as its release) when followed within the syllable by another vowel. The retroflex phoneme /r/ occurs simultaneously with a preceding /h, t, th/ or /u, e, a/; in the environment e, a....r, the nasal phoneme /ng/ is simultaneous with the preceding low or mid vowel and the following retroflex vowel, with the result a nasalized retroflex vowel. With many younger-generation speakers, however, the sequence /ngr/ is lacking; such speakers make kɑngrl 'small jar' sound like kɑrl 'liver'.

12. **Conclusions.** We have found that the pairs of phonemes /e:/ and /e/, /ɑ/ and /a/ (and possible /ɑ:/), and /5/ and /3/, are in each instance distinctive by virtue of certain limited morphophonemic operations. A subsidiary system is clearly in evidence which covers all the data except that brought about by alternations in certain morpheme complexes. A second, fragmentary system consisting of three (or possibly four) additional phonemes

must be split off to account for this divergent data; the two systems integrated together constitute the total phonemic structure of the dialect.

For certain purposes, such as morphophonemic notation, historical and comparative work, lexicography, and as a limited illustration of phonemic principles, it is certainly convenient to use the primary sub-system, in which /e:/ is in complementary distribution with /e/, /ɑ/ with /a/, and /5/ with /3/. We could introduce into our system of notation a morphophonemic disturbant (using for example the colon) to account for divergent phenomena with the phoneme /e/ (and maybe /ɑ/) to be read 'long allophone where the short would be expected from the environment'. For the front and back low vowels, it is enough to write morphophonemically (than1 pu0....and tham1 pu0....) and treat 'we (inclusive)' as tzanm. In a similar vein, we can write morphophonemically the neutral-tone syllables of Tone-3 origin; the lack of stress automatically indicates the neutral tone (to my ear, though apparently not to Hockett's), but the converse is not quite true, as pointed out above, because of interjections.

In this way, we could simplify our notation for the overwhelming mass of data, by writing the few conflicting instances noted above as follows: IIγue:m3 me0 'us??' and IIγuem3 me0 'steady?' (or, morphophonemically, IIγuen3 me0) but IIγue3 'me' (where there is no possibility the vowel could be short); IIke:rl 'song' and IIkerl 'root'; IImai2 tiar3 'bury a little' and IImai3 tiar3 'buy a little' but IImai3 te0 'bought'; IIhiau3 kie3 'young lady' but IIkie3 kie0 'elder sister'. The distinction between IItre4 hr0 'this is' and IItre:4 hr0 'this is (sarcastic)' can either be ignored as stylistic or noted by writing the former IItre4-hr0, utilizing a sort of purely graphic juncture-device.

It certainly proves convenient not to write the phoneme /γ/, instead separating the syllables by some device (e. g. the tone markers), and making a morphological statement about the linking particles. It is unnecessary, too, to write the neutral-tone phoneme /0/ when it occurs (as is usual) with an unstressed syllable, since (to my ear) the lack of stress is always accompanied by /0/. These various devices of convenience should not compromise the nature of the underlying analysis, of course.

I hope that some of the notions pursued here will prove handy in other areas. Cantonese, for example, has a colloquial morpheme which corresponds to the Peking retroflex suffix in meaning; it appears in the shape of two tone

changes, one of which introduces a new tone phoneme, while the other disturbs the existing complementation of the so-called "entering" tone with other tones, and also the complementation of the two allophones of the so-called "upper-even" tone. These disturbances are not apparent in the reading pronunciation (a more wide-spread style in China. it should be noted, than in the West), but only in conversational speech.[11] In the Changsha dialect, similarly, an additional tone is required by colloquial speech as contrasted with reading speech; a lower-register variety of the so-called "going" tone. The concept of subsidiary phonemic systems may also provide an answer to the interesting theoretical questions posed by intersecting vowel allophones in the complicated tonal sandhi of the Foochow dialect.

We have found it possible to analyze the affricates and spirants of Mandarin in a way drastically different from that of Hartman and Hockett, and have suggested that off-hand notions of "phonetic similarity" do not provide a determinate answer to the problem of choosing between multiple phonemicizations; it has yet to be proved that the ultimate phonologic constituents of an idiolect can ever be uniquely identified.

11. Cf. Chao, Cantonese Primer 34-6 (Cambridge 1947).