An experimental investigation into phonetic symbolism as it relates to Mandarin Chinese[1]

RANDY J. LAPOLLA

10.1.   Introduction

10.1.1.   General

About the time Plato was writing *Cratylus*, in which the protagonist argues that the relationship of sign to signified was one of object to imitation (Fowler 1977), that is, less than arbitrary, discussion of this question was also going on in China, especially among the Confucianists. The Chinese, in general, came to a different conclusion than Plato's Cratylus. The writer who spoke most directly to this question was Xun Zi (d. 221 BC) (Hong 1982). He felt that an object and its name had a totally arbitrary relationship, that no name was any more "suitable" than any other, that all names are "suitable" only as a result of convention and popular usage. Though there were some Chinese philosophers (from the Han Dynasty on) who attempted to find the *true* sound–meaning correspondences for words, going against Xun Zi's principle of non-suitability of names, the mainstream of Chinese linguistics to this day still does not really question the concept of total arbitrariness. Almost all of the general books on linguistics now in use in the People's Republic (e.g. Gao and Shi 1963; Ye and Xu 1981; Ma 1981) use the same quote from Xun Zi and another, from *Capital*, where Karl Marx says "The name of a thing is entirely external to its nature" (Fowkes 1977: 195).

In the West, although Saussure felt that "No one disputes the principle of the arbitrary nature of the [linguistic] sign" (1966: 68), there have been many since that time who have tried to show that the assignment of signifier to signified is not always completely arbitrary. Jespersen (1922) went to great lengths to show that among the languages he was familiar with there were definite correlations between certain types of sounds and certain categories of meaning, such as size, movement, feelings, and distances. Jespersen also felt that people have an instinctive feeling for what are the best sounds for the different meanings in each category.

10.1.2.   Experimentation

Much of the debate between those who support the idea of sound symbolism and those who do not, and between those who support its universality and those who do not, has centered around methodology of experimentation.

Sapir (1929) was the first to try to prove the latter of Jespersen's hypotheses experimentally (particularly a feeling for "large" and "small" among English speakers), and was followed by Newman (1933), who tried to support Sapir's findings and expand on them by including the feeling for "bright" and "dark," and by examining the lexicon for sound–meaning correspondences. Though these two studies came under some criticism, the conclusions they felt their results led to set the tone for much of the discussion of sound symbolism that has taken place since that time. The five conclusions they drew were that

(1) Phonetic elements tend to be rigidly patterned on a non-linguistic symbolic scale . . . (2) The factor of age . . . has little effect on the subjective scale . . . (3) The basis of phonetic symbolism is fundamentally objective . . . (4) Diverse types of phonetic patterns are formed by unlike symbolisms[2] . . . (5) These symbolic judgements are not produced by linguistic associations (Newman 1933: 75).

Since that time there have been some experimental investigations that have claimed to have disproved the theory of sound symbolism (Bentley and Varon 1933; Maltzman *et al.* 1956; Brackbill and Little 1957), but these results can often be explained as being due to important procedural problems (Brown and Nuttall 1959; Weiss 1963).

Starting with Brown *et al.* in 1955, a ten-year debate on the pros and cons of various types of methodology of experimentation and evaluation of results ran through the *Journal of Abnormal and Social Psychology* and later the *Psychological Bulletin* (Brackbill and Little; Brown and Nuttall 1959; Miron 1961; Weiss 1963; Taylor 1963; Weiss 1964; Johnson *et al.* 1964; Taylor and Taylor 1965), but no agreement on the best method was reached.

At least ten experiments have been done on sound symbolism using either the Chinese language (Mandarin) or Chinese subjects. (Sapir 1929;[3] Brown *et al.* 1955; Brackbill and Little 1957; Brown and Nuttall 1959; Weiss 1963; Huang *et al.* 1968; Tsien-Lee 1969; Klank *et al.* 1971; Lester 1973, 1974.) Of these, nine had results supporting the sound-symbolism hypothesis, and only one (Brackbill and Little 1957) had results which did not support the hypothesis. With all of these experiments I believe there were problems of methodology and procedure serious enough to warrant a new set of experiments.

10.1.3   Tonal morphology

10.1.3.1   Monosyllabic

A distinctive feature of a large group of languages, including many Southeast Asian languages and several Chinese dialects in the southeastern part of China, is the use of tone change to mark morphological or semantic change. Though this use of tone is not always limited to marking the hypocoristic or diminutive,[4] this is one of its more common uses. This is true of Lahu, Yi (Lolo),[5] and the Hui, Wu (that of southern Zhejiang), and Yue dialect groups of Chinese (Hirata 1983). In the Canton dialect (one of the Yue dialects), for example, which is one of the most representative and best attested of the Chinese dialects using tone change to mark the hypocoristic, a change to a high level or high rising tone is used to mark the diminutive or familiar.[6]

In standard Mandarin Chinese the hypocoristic of nouns is marked only with the affixation of a retroflex suffix,[7] with no change in the original tone. Even so, if there is a universal tendency to associate high frequency with "smallness" (the diminutive in general), then Mandarin speakers also should be sensitive to tone differences such as those in Cantonese.[8]

10.1.3.2.   Disyllabic

A second use of the change to a high tone in Cantonese is to change the intensity of reduplicated adjectives. This can be done to different degrees depending on which syllable the changed tone is assigned to. Lengthening of the vowel is also used for extreme intensification. For example, [hong^{21}] "red" can be modified to [ho:ng^{35} hong^{21}] "very red" or [hong^{21} hong^{35} tei^{35}] "a little red."[9] (Yue-Hashimoto 1972: 95; Whitaker 1955–1956: 31) The first type of tone change is also obligatory on the first syllable of a three-syllable adjective formed by reduplication of the first syllable of a disyllabic adjective for the purpose of intensification, as in [søn^{35} søn^{21} pan^{35}] "very good tempered" from [søn^{21} pan^{35}] "good tempered." The second type of tone change is possible for onomatopoeic (either in the metaphysical or real sense) phrases with an initial reduplicated syllable, as in [lap^{33} lap^{35} lyn^{33}] "disorderly"[10] (Yue-Hasimoto 1972: 95).

The change of the first syllable to a high tone is obligatory when reduplication is used for extreme intensification, and can apply to syllables in all of the nine tones of Cantonese because of the lengthening of the syllables. The tone change in the second type of reduplication, though, is optional, and can only apply to syllables with low tones (Yue-Hashimoto 1972: 99). Because with the tone change the resulting tone pattern (i.e. low-high) is the same as that for familiarity, it is possible that the changed form is less intense or serious, as with the phrase [t'an^{21} t'an^{35} tsan^{44}] "shivering," because the change puts it in the field of the hypocoristic. It is

also the case, in Thai and the Hainan form of the Southern Min dialect of Chinese, that modification of the first syllable of a reduplicated adjective by raising the tone and often lengthening the vowel is used for the purpose of intensification. In Thai, there is [dii^{55} dii^{33}] "very good," from [dii^{33}] "good" (Li 1977: 8; Haas 1946: 128–30), and in Hainan Chinese there is [bui^{55} bui^{33}] "very fat" from [bui^{33}] "fat" (Woon 1979: 87). According to Lien Chinfa (personal communication), this is also the case in mainland Southern Min.

What is significant in all these languages is that for extreme intensification, modification must be of the first syllable and the resulting tone pattern is high-low.[11] In Cantonese and in Thai, other types of adverbial compounds or marked disyllabic words will be marked by modification of the second syllable, so generally have a low-high tone pattern. In Cantonese this is also the case for marking familiarity, especially family relation terms, such as [pa^{22} pa^{55}] "daddy" or [mui^{22} mui^{35}] "sis" (Gao 1980: 22–23).

In Mandarin, the tone on the second token of a reduplicated adjective changes to a high level (55) tone, but this change of tone is not usually thought to be semantically significant. Although this is probably the case, I wanted to see if native Mandarin speakers would also be sensitive to the tone pattern-to-meaning correspondences in Cantonese.

10.1.4.   The hypotheses

The hypotheses we attempted to test in this investigation are similar to those tested by other researchers: (1) that English speakers should have better-than-chance-expectancy success at matching Chinese antonymic word pairs, thereby showing that Chinese exhibits sound-symbolic patterns which are perceptible to English monolingual speakers, and (2) that Chinese native speakers would have enough of a common innate sense of sound–meaning correspondences to give a common pattern of responses when asked to assign meanings to sounds.

10.1.5.   General methodology

The method we used in the first experiment of the investigation was one of those that is generally most supported, matching antonymic pairs, being careful to eliminate or reduce any of those factors that have been called into question.[12] But for one part of the investigation using Chinese subjects (Experiment 2, Part a), we used a new method that gave the subjects much more freedom in making sound–meaning choices, while at the same time allowing us to analyze the choices into grave and non-grave segments, a key parameter of sound symbolism (Jakobson 1978).

10.2. Experiment 1

10.2.1. Method

10.2.1.1. Part a

The first part of the first experiment was designed to test the ability of native English speakers who had never been exposed to Chinese to match the correct word out of a pair of antonymic Chinese words and the written English translation of one of them. A significant number of correct answers would point to the existence of sound-symbolic patterns in the Chinese lexicon, and would also be evidence that English speakers are sensitive to them.

A list of 40 English words that are members of common antonymic pairs was given to each of the subjects.[13] Beside each word were the numbers 1 and 2, and the word "neither."[14] A tape recording of a list of antonymic pairs that were the Chinese equivalents of each English word and its commonly accepted antonym, spoken by a native speaker of Mandarin from Taipei,[15] was played for the subjects. Each pair was spoken with two seconds between the members of the pair, and there was a five-second pause between pairs. The subjects were asked to circle 1 or 2, depending upon which word of each antonymic Chinese pair they thought had the same meaning as the English word given, or "neither," if they felt that neither of the words "felt right." They were told that some pairs might be similar in meaning to other pairs, but to judge each pair separately. The word list used, and the Chinese words[16] spoken for each item, are given in table 10.1.

The Chinese words were spoken because giving a romanized transliteration of the Chinese words would not give linguistically naive subjects a good idea of what the sound was really like, especially the tone. This could also lead to the subjects matching the words because of orthographic similarities.

Only one English word was given for each pair, for three reasons. First, it simplified the answering process for the subject, comparing one word with two instead of two with two. Second, it reduced the possibility of the subject just matching up the vowels of one or both of the English words with that/those of the Chinese words, or basing his/her choice on the English–Chinese pair that was not being asked for.[17] Third, it allowed us to use the other half of some of the English pairs later on in the test, with the same Chinese pair, either in the same order or not, depending on randomization. This third point was analyzed separately, and used as a kind of double check. Since the subjects were told that many pairs were possibly somewhat synonymous, they might not have noticed the same pair being given twice. If they did not notice, but got the correct answer on both, then it would reinforce the findings of the experiment; if they did not notice and had opposite results for the two times a pair was used, then any positive results would be cast in doubt. To find out if the subject noticed or not, the question was asked at

Table 10.1 *The antonymic pairs tested*

ugly	měi	chǒu	cold	lěng	rè
long	dǔan	cháng	shallow	shēn	qiǎn
many	dūo	shǎo	concentrated	dàn	nóng
thick	báo	hòu	micro	wēi	jù
big	dà	xiǎo	friend	dí	yǒu
sharp	rùi	dùn	few	shǎo	dūo
bright	liàng	àn	small	xiǎo	dà
coarse	xì	cū	narrow	kuān	zhǎi
fat	pàng	shòu	hard	yìng	ruǎn
soft	ruǎn	yìng	skinny	shòu	pàng
heavy	zhòng	qīng	macro	jù	wēi
weak	qiáng	ruò	strong	ruò	qiáng
wide	zhǎi	kuān	hot	rè	lěng
tall	gāo	ǎi	diluted	nóng	dàn
fast	kuài	màn	thin	hòu	báo
drunk	xǐng	zùi	short	ǎi	gāo
joy	xǐ	bēi	high	gāo	dí
dry	gān	shī	fine	cū	xì
low	dí	gāo	slow	màn	kuài
quiet	xuān	jìng	sorrow	bēi	xǐ

the end of the test: "Did you feel any of the separate pairs you heard on the tape was the same as any other pair, though possibly not with the same internal order? If so, which ones?"

10.2.1.2. Part b

For the second part of the first experiment, a second group of ten native English speakers was given the same test as in Part a, except that a second recording, using a different native Mandarin speaker from Taiwan,[18] was made in which the tones were reversed on the words in those pairs whose members differed in tone. This was done to see if the tones of the words could have been instrumental in the choices made by the first group of subjects.[19] This is of course assuming that the second group of English speakers would have roughly the same results as the first group if given the exact same test.

10.2.2. Results

The number of right and wrong answers[20] from Part a out of the total non-"neither" responses were calculated and are presented in line 1 of table 10.2. The

Table 10.2 *Results of Experiment 1*

Type	No. of responses	Right	Wrong	
Part a	359	190	169	$(p < 0.30)$
Part b	337	187	150	$(p < 0.05)$

Table 10.3 *Comparison of items from Parts a and b*

word	Part a	Part b	Part a result	Part b result
big	dà	dǎ	9/0	4/4
coarse	cū	cù	6/2	9/0
soft	ruǎn	ruàn	9/1	2/6
wide	kuān	kuǎn	1/8	6/1

subjects in general only noticed one or two repetitions, and were not sure which ones had been repeated, so it seems this did not interfere with the results. The number of correct answers was approximately the same for both halves of those antonymic pairs that were split up and tested separately.

The results for Part b, calculated the same way as for Part a, are given in line 2 of Table 10.2.

These results tell us that tone could have been an important criterion for the subjects' judgments.[21] It seems here that the change in tone made it easier for the subjects to pick the correct Chinese word. This could mean either that the tones assigned to some words are not the ideal ones from an English speaker's point of view, or that something about the difference between the two recordings or the two groups of subjects influenced the results.

For some pairs the difference before and after the tone switch was quite startling. Table 10.3 contains some of the pairs that experienced the biggest change in response. The number to the left of the slash in each column is the number of right responses, and to the right is the number of wrong responses.

Some of the differences here match up with the results from Part a of Experiment 2 (see 10.3.2.1): as the high level tone is favored for the "small" category, the better scores for *coarse* and *wide* follow logically. The results for *big* also follow logically from our results showing a preference for the falling tone for the "big" category, but for *soft*, the effect is the opposite. The better score for *soft* before the change to a falling tone seems to contradict the tendency for the use of the falling tone for "big" category words, unless the "big" category is not a monolith, and can

be subdivided into several different semantic subgroupings, such as "soft," "warm," and "big." In this case, it seems the falling tone might be less suitable for those words connoting softness.

10.3. Experiment 2

10.3.1. Method

10.3.1.1. Part a

The second experiment was designed to test the subjective sensitivity of Chinese speakers to sound symbolism. In Part a, a list of 48 English words that are all members of antonymic evaluative pairs[22] was printed on the answer sheets, with a blank line next to each word. A separate group of 50 nonsense syllables was taken from a list of all possible Chinese syllable types (from Zhong 1980) and printed in five columns of ten words each. They were chosen because they were simple CV syllables, with the vowels all monophthongs and the consonants all stops or fricatives, all of which could be easily classified as [± nasal] and/or [± grave]. No recording was made of the words.

The subjects were five native Mandarin-speaking graduate students from Taiwan. They were instructed: "Please say these words to yourself, then assign these words to the meanings given on the answer sheet, based only on how appropriate you feel they sound for the given meaning. Also assign tones to the words, based on the same criterion."

10.3.1.2. Part b

For Part b, a list of 25 Cantonese words or phrases was compiled from Chao 1947, Whitaker 1955–1956, Yue-Hashimoto 1972, and Kam 1977, from their lists of words and phrases that can be either made diminutive/familiar or, in the case of adjectives and adverbs, changed with regard to intensity by a change in the pronunciation to a high- or high-rising tone. Questions were then formed to contrast the semantic differences between the minimal pairs. Each question was then recorded (spoken by myself), followed by the relevant minimal pair, the order having been randomized, spoken by a native speaker of Cantonese from Hong Kong.[23] The subjects, all native Mandarin speakers from Beijing or Taiwan who had no knowledge of Cantonese, were told that they would hear a tape of Cantonese words, and that they should answer only "A" or "B" after hearing each pair, based on which one they thought was the best answer to the question asked.

10.3.2. Results

10.3.2.1. Part a

The 48 English words used for this part of the experiment were divided into "small" category words and "big" category words. Out of the total 210 choices (initial and vowel) made by the subjects for the "big" category words, grave segments were chosen 151 times, much greater than chance probability ($p < 0.001$), but for the "small" category words, grave segments were chosen only 103 times, no better than chance probability. The number of grave initials and grave vowels chosen by the subjects for each category are given in table 10.4.

The results show a clear tendency on the part of the subjects to assign grave consonants and vowels to "big" category meanings, and not to the "small" category meanings. This agrees with the findings of Solomon 1959: 494, which showed that subjects' judgments of "heavy" or "high magnitude" relate to the lower octave bands, while judgments of "light" or "low magnitude" relate to the higher octave bands. It also correlates well with the findings of Huang et al. 1969, which showed that when asked to produce words meaning "large" or "small," subjects produced words that overwhelmingly favored the acute vowels for "small" and grave vowels for "large."

Two other measurements were done, one on the nasals and one on the tones. A count of the number and type of nasal used for each category is given in table 10.5.

From this table we can see that the subjects made a clear choice of the grave nasal for the "big" category, and the acute nasal for the "small" category.

The number of times each tone was used for each category is given in table 10.6. The numbers 1, 2, 3, 4, and 5 refer to the high level, high-rising, falling-rising, falling, and neutral tones respectively.

We can see at least two tendencies here. One is the preferred use of the high level tone for the "small" category meanings, and the other is the preferred use of the falling tone for the "big" category meanings.

It is interesting to note that the tendency on the part of the subjects to use the high-level and falling tones more than the other tones correlates well with the findings of Yue-Hashimoto 1980 and Li and Thompson 1977. These studies showed that the high-level and the falling are the first two tones to be mastered by a child learning to speak Mandarin as its native language. Tse (1977) has reported that for Cantonese speakers as well, the high-level tone is the first to be acquired. It would be hard to determine if this is because this tone is easier to learn and produce (the "Difficulty Hypothesis"[24] of Li and Thompson), or because it is compatible with the size and defenselessness of the child (see discussion of "frequency code" below).

Table 10.4 *Analysis of grave segments*

Type	No. of responses	Grave initial	Grave vowel
big	105	71	80
small	105	52	51

Table 10.5 *Comparison of nasals*

Type	No. of responses	No. of [m]	No. of [n]	
big	23	17	6	($p < 0.01$)
small	29	9	20	($p < 0.01$)

Table 10.6 *Results for each tone-type*

	N	1(55)	2(35)	3(214)	4(51)	5(var.)	
big	105	13.3	16.2	15.2	50.5	4.8	($p < 0.001$)
small	105	35.2	16.2	16.2	20.9	11.4	($p < 0.01$)

10.3.2.2. Part b

To analyze the results of this part of the experiment, the questions were divided into four types, depending on what was being tested: (a) the hypocoristic (i.e. size-to-sound and monosyllabic familiarity); (b) "onomatopoeic" intensification; (c) disyllabic familiarity; and (d) intensification of reduplicated adverbs. Out of the 150 responses to type (a) questions, 87 were the same as they would be for a native Cantonese speaker (the definition of "correct") ($p = 0.05$). If we break this down into the subtypes, we see that of the size-to-sound responses, only 51 out of 100 were correct. Of the 50 responses to the questions testing monosyllabic familiarity, 36 were correct ($p < 0.01$).

Fifteen responses out of the twenty for type (b) pointed to a feeling on the part of the subjects that the unmodified form was the more intense of the two ($p < 0.05$). As I pointed out in section 10.1.3.2, this is the same for some native Cantonese speakers. Twenty-one of the 40 responses were correct for type (c).

The results for type (d) were much more straightforward and strongly support the universality of the feeling for the high-low pattern as an extreme intensive. Thirty-four of the 40 responses were correct ($p < 0.001$).

10.4. Discussion

I would like to explain these results by reference to the "frequency code" theory as developed in Ohala 1982a, 1982b, 1982c, 1983, 1986, and Morton 1977. As mentioned in the introduction of this volume, "frequency code" refers to the universal communication of size by frequency between vocalizing animals in close-contact competitive encounters. In such an encounter, an aggressor would try to make itself seem as large as possible to intimidate its opponent into submission. If this were successful, the opponent would try to make itself seem as small and unthreatening as possible, sometimes even acting infantile, to avoid being hurt. One of the ways for both sides to achieve their different ends would be by altering the frequency of their vocalizations. By using low-frequency, staccato vocalizations, the aggressor makes himself seem larger and more threatening, as generally the larger any object or animal is, the lower and more irregular the sound it will emit. In the same way, an animal that wants to capitulate needs to appear non-threatening, so would use high-frequency, tone-like vocalizations to appear smaller or infant-like. This code is not something that is learned, but is innate (Morton 1977, 1986).

In human language there is a universal use of similar pitch contours for statements and questions that are not lexically or syntactically marked (Bolinger 1964, 1978). Many languages also use high pitch for polite speech (Brown and Levinson 1987: 267–268). Just as an animal seeks to mollify its opponent by using a high-pitch tone-like cry, it is possible that human languages use a high or rising pitch for questions and polite speech because in asking a question, or in being polite, the speaker defers to the addressee, in a sense admitting that the other party is stronger, at least as regards the situation or topic of conversation. The other part's answer is in a falling-pitch pattern, again to reflect this power relationship. Friendliness is also shown by using high-frequency (especially low-high) patterns more, to show that the speaker is not a threat to the other party.

We have seen from the results of these experiments that there is a cross-linguistic tendency toward associating acute segments with "small" category words, and grave segments with "big" category words. As acute sounds have energy largely in the high frequencies, and grave sounds mostly in the low, this could be related to the "small" ~ high frequency and "big" ~ low frequency relationships mentioned above. Just as objects and animals emit sounds of different frequency corresponding to their sizes, it seems natural that people, by extension, would name these things, or adjectives of size, etc., according to the same criterion.

This theory does much to explain some heretofore unexplainable facts of nature and language, such as the origin of the smile and some other facial features (Ohala 1980), the reason for sexual dimorphism in the vocal anatomy of many species of vocalizing animals and birds (Ohala 1983, 1986), and size–sound symbolism in language (Ohala 1982a, 1982b, 1982c, 1983, 1986). This could also help explain the

more frequent use of rising intonation in "women's speech" (Lakoff 1975: 17) and why, as Whitaker (1955–1956) and Haas (1946: 130) have pointed out for Cantonese and Thai respectively, women are more likely than men to use the changed tone patterns in optional cases. The fact that the changed tone form of Cantonese is used more frequently in informal speech (Wong 1982) might also be explained by reference to this.

This theory does not explain everything to do with sound symbolism, though, such as why *soft* is usually grouped with the "big" category words. It may be (as mentioned in section 10.2.2) that the "big" category is not a monolith, but is comprised of smaller categories that overlap in their use of sounds, even if the source of the sound–meaning correspondence is very different. This would be the case, for example, if *soft* became associated with grave sounds because the sound of a mother's heartbeat that a fetus or infant hears is a low non-tone-like sound.[25] For many vocalizing animals and birds, the mother's call to the young is a low soft sound. For some of these creatures there is a change in voice quality (to soft) and/or a drop in the pitch of the voice of the female that accompanies the birth of the young (Collias 1960).

The relationship of amplitude and pitch, and other related questions, could be solved with enough research. It would be interesting to do further experimentation to see if, as reported in Hata 1983, some phonetic features have tighter sound–meaning correlations than others, and also to repeat these experiments using color terms (cf. Solomon 1959; Fischer-Jørgensen 1978).

NOTES

1 I would like to thank Leanne Hinton, Robert Blust, John J. Ohala and Johanna Nichols for valuable comments on earlier versions of this paper, and Jing Fang Wang for her help on the experiments.

2 Points (3) and (4) are based on the experimenters' attempt to explain their findings using either the kinesthetic factors of articulation, acoustics, or a combination of the two.

3 Sapir used seven Chinese subjects for his experiments, but did not give any more information about them (such as age, native dialect, time in US, etc.) other than the fact that they were "Chinese." (Li Fang-Kuei, personal communication, has informed me that he was one of the seven Chinese subjects.) He gives the age and status, but not the race or nationality, of his other subjects. The results for the seven Chinese subjects were not included in the general results, but were reported to be similar to those for the other subjects.

4 For a discussion of the different uses of tone change in Cantonese other than for the hypocoristic, see Kam 1977 and 1980.

5 In Lahu there are two parts to making the diminutive-extensive: affixation of a non-low unrounded front vowel, and a change of tone from the mid level (33) (tones here are

marked with the five-point scale first introduced in Chao 1930) tone of the neutral extensive to the high-rising (45) tone of the diminutive-extensive, e.g.

[chi³³ hi³³] "this size" – [chi³³ hii⁴⁵] "such a small one" (Matisoff 1973: 18).

In Yi, there is a raising of tone and a devoicing of some vowels to mark the hypocoristic, e.g.

[a¹¹ mæ³³ zu¹¹] "a girl" (you do not particularly like) – [a¹¹ mæ⁵⁵ zu³³] "a small girl" (you like). (Ma Xueliang, personal communication.)

6 Cf. Solomon 1959: 496, which showed that for the subjects of that experiment, "sounds with energy concentrated in the lower octave bands are reliably judged as 'strange' . . . while sounds with energy concentrated in the higher octave bands are judged 'familiar'."

7 For discussion of the relationship of the retroflex final to the changed tone and the possible source of the changed tone, see Kam 1980; Chao 1945; Hirata 1983; Whitaker 1955–1956; and Wong 1982.

8 It is interesting to note the differences in the tone systems of the two languages. In Cantonese, there are three pitch levels that are significant, and most of the tones are in the lower two levels. Gandour (1981) has shown that contour and direction, in general, are more salient features than height for native speakers' distinguishing of the different tones in Cantonese. This is contrary to the findings of Vance 1976, but is logical given the closeness in register of the four non-stopped low tones. For showing markedness, though, the most salient feature of the changed tones is the change from [− high] to [+ high]. This correlates well with two of Maddieson's universals of tone: "3. Phonetically Central tones are unmarked, Extreme tones are highly marked . . . 4. Systems in which high tones are marked are more frequent than systems in which low tones are marked" (Maddieson 1978: 341–342). This is reflected also in the fact that the low tones are statistically more likely to undergo this change (Wong 1982). In Mandarin, though it is possible to see the third tone as a low tone (Hashimoto 1981), register is not very important, possibly because all of the tones can at some point in their production be in the [+ high] register. This is supported by the findings of Victor Zue (reported in Klatt 1973), where even when compressed into a four-kHz range the tones on synthetic Mandarin syllables were judged correctly by the subjects 90% of the time. Even when compressed to two kHz the level and the falling tone were still judged correctly, while the other two tones were confused. Wu (1984) also showed that register was not necessary in an autosegmental analysis of Mandarin tones.

9 According to Samuel H-N. Cheung (personal communication), there is another type of modification of some adjectives using reduplication of an onomatopoeic syllable after the adjective. With this type of modification, a reduplicated high-tone onomatopoeic syllable will usually give the adjective positive connotations, and a low-tone syllable will give the adjective negative connotations, as in [fei²¹ tyt⁵⁵ tyt⁵⁵] "plump" or "fat (like a nice fat baby)" and [fei²¹ than²¹ than²¹] "fat" (as a derogatory reference to fat people).

10 According to Whitaker (1955–1956), there is an intensified form of this phrase, with the tone change on the first syllable, but it is not a reduplicated form: [lap³⁵ kam³³ lyn³³] "very disorderly."

11 It could be argued that these languages are all related, but as Haas (1946: 130) points out, high-low tone is also used by some speakers of English. She gives the example of

Louisiana, where adjectives such as "good" can be intensified by reduplication where the first syllable is lengthened and its tone is raised, as in *goo-ood good*. In English the low-high pattern for familiarity (especially when talking to babies) is also very common.

12 See Gebels 1969 for a discussion of some of the possible problems associated with this methodology.

13 The words were actually 25 pairs of antonyms split up, with redemption of 15 out of the 25 semantic fields represented by the 25 pairs. The words were chosen from Brown *et al.* 1955, and Lu 1981. Some of the pairs Brown used, such as "many ∼ one," are not normal antonymic pairs, and so were not used.

14 The subjects were given the choice "neither," so that they would not be forced into making choices they were not comfortable with. As even proponents of the theory of phonetic symbolism do not claim that all words are symbolic, I feel allowing the subjects this option is necessary. After running these experiments I came across an interesting study by Asher Koriat (1975) that took the subjects' "feeling of knowing" into account by asking them to rate how certain they felt about each choice they made. The findings show that there was an awareness on the part of the subjects of phonetic symbolism, so the choices with the higher confidence scores were also the ones judged correctly most often. It is suggested that these higher scores are the more relevant ones. If I were to do these experiments over, I would also use the confidence-rating test instead of giving the "neither" choice.

15 I do not feel that having a native speaker (not connected with the experiment) say the words could possibly undermine the results of the experiment. If the native speaker (unaware of the purpose of the experiment) emphasizes those aspects of the speech sounds s/he feels most crucial to correct perception of the meaning of the word, such as those mentioned by Jespersen (1922), and the subjects respond to the same parameters, then this is further proof of the universality and direction of phonetic symbolism. It would be possible to do a test just asking speakers to emphasize or de-emphasize certain words, to see what sounds they change and in what way. The recreation in English of *tini* [tini] after the Great Vowel Shift changed *tiny* [tini] to [taini], and the creation of words like *humongous* from *huge* probably reflect this tendency to emphasize certain sounds for certain meanings. It would also be interesting to see, if sound changes diffuse across a lexicon, as discussed by Chen and Wang (1974) and Wang (1979), whether or not sound-symbolic words would be more resistant than non-sound-symbolic words to any change that would make them less sound-symbolic. This might partially explain why some words would undergo a certain sound change earlier than others.

16 The Chinese is given in pinyin romanization. The tone symbols are ā high level, á rising, ǎ falling-rising, and à falling.

17 For example, if I wanted the meaning of the italicized word in a pair, such as "*beautiful* ∼ ugly" Chinese "meǐ ∼ chòu," I might get the right answer not because the subject felt there was a relationship between "beautiful" and "meǐ," but because he felt one between "ugly" and "chou." This could still happen using our methodology, but we are attempting to reduce the likelihood of it happening.

18 Ideally, the original native speaker should have been used for both of the recordings, but this was not possible.

19 Given enough time and subjects, this second part could have been repeated two or three more times, each time reversing a different set of segments (such as the initials) to see

which segments and which position in the syllable structure has the greatest effect on the subjects' choice.

20 Here "right" represents the subject's picking the correct segmental form with the altered tone, and "wrong" represents the opposite case.

21 This contradicts the findings of Lester 1974, though as Lester's test presented the subjects with English words and Chinese tone marks written on paper, I do not feel the results can be valid.

22 The list was the same as that for the base list in Experiment 1, but without the pair "wet ~ dry" (left out because I wasn't sure how to analyse these words in terms of "big" ~ "small"). Fifty Chinese syllables were used because this was the number of syllables in Chinese that fit the criteria. English was used, and not Chinese characters, to diminish the chance that the subjects would match up sound and character.

23 There is no difference in this aspect between the Canton dialect and the Hong Kong dialect (Kam 1977; Cheung 1969).

24 If the "Difficulty Hypothesis" is valid, the use of the high-rising tone for markedness could possibly be explained by reference to the relative difficulty of producing it, as a tone that is more difficult to produce might seem less natural, and so therefore marked. As Gandour (1977: 60) has reported that it is natural for syllables with low tones to be longer than syllables with high tones, the lengthening of the initial syllable that accompanies the change in tone for extreme intensification of reduplicated adjectives might also be explained by reference to ease or naturalness of production and markedness.

25 Brown and Levinson (1987: 268) discuss creaky voice as a feature of "positive politeness" (as opposed to the deferential use of high pitch, which is a feature of "negative politeness"): "[C]reaky voice, having as a natural source low speech energy, can implicate calmness and assurance and thence comfort and commiseration . . ."

REFERENCES

Bentley, M. and E. J. Varon. 1933. An accessory study of phonetic symbolism. *American Journal of Psychology* 45: 76–86.

Bolinger, D. L. 1964. Intonation as a universal. In H. G. Lunt (ed.) *Proceedings of the Ninth International Congress of Linguists, 1962*. The Hague: Mouton, 833–844.

1978. Intonation across languages. In J. H. Greenberg (ed.) *Universals of Human Language II: Phonology*. Stanford: University Press, 471–524.

Brackbill, Y. and K. Little. 1957. Factors determining the guessing of meanings of foreign words. *The Journal of Abnormal and Social Psychology* 54: 312–318.

Brown, P. and S. C. Levinson. 1987. *Politeness: Some Universals in Language Use*. Cambridge: University Press.

Brown, R. W., A. H. Black, and A. E. Horowitz. 1955. Phonetic symbolism in natural languages. *The Journal of Abnormal and Social Psychology* 50: 388–393.

Brown, R. W. and R. Nuttall. 1959. Methods in phonetic symbolism experiments. *The Journal of Abnormal and Social Psychology* 59: 441–445.

Chao, Y. R. 1930. A System of tone letters. *Le Maître Phonétique* 45: 24–27.

1945. The morphemic status of certain Chinese tones. *The Transaction of the International Conference of Orientalists in Japan* 4: 44–48.

Chen, M. Y. and Wang, W. S.-Y. 1974. Sound change: actuation and implementation. Paper presented at First International Conference on Historical Linguistics, Edinburgh, September 1973. Revised March 1974.

Cheung, Y.-S. 1969. A study on the upper even tone and changed tones in Cantonese as spoken in Hong Kong. *Hong Kong Chinese University Chinese Culture Research Institute Journal* 2: 81–107. (In Chinese.)

Collias, N. E. 1960. An ecological and functional classification of animal sounds. In W. E. Lanyon and W. N. Tavolga (eds.) *Animal Sounds and Communication*. Washington DC: American Institute of the Biological Sciences. Publication no. 7, 368–391.

Fischer-Jørgensen, E. 1978. On the universal character of phonetic symbolism with special reference to vowels. *Studia Linguistica* 32: 80–90.

Fowkes, B. (trans.) 1977. Karl Marx, *Capital: A Critique of Political Economy*, vol. I. New York: Random House.

Fowler, H. N. (trans.) 1977. *Plato in Twelve Volumes, 4: Cratylus, Parmenides, Greater Hippias, Lesser Hippias*. London: William Heinemann Ltd.

Gandour, J. 1977. Interaction between tone and vowel length. *Phonetica* 34(1): 54–65.

1981. Perceptual dimensions of tone: evidence from Cantonese. *Journal of Chinese Linguistics* 9(1): 20–36.

Gao, H. 1980. "Research on the Canton Dialect." Hong Kong: Commercial Press. (In Chinese.)

Gao, M. and A. Shi. 1963. "A Survey of Linguistics." Beijing: Zhonghua Shuju. (In Chinese.)

Gebels, G. 1969. An investigation of phonetic symbolism in different cultures. *Journal of Verbal Learning and Verbal Behavior* 8: 310–312.

Haas, M. R. 1946. Techniques of intensifying in Thai. *Word* 2: 127–130.

Hashimoto, M. J. 1981. A phonological characterization of syllabic intonations in the so-called tone languages. In A. B. Gonzalez and D. Thomas (eds.) *Richard S. Pittman Festschrift*. Manila: De La Salle University Press.

Hata, K. 1983. Two experiments in sound symbolism. ms. University of California at Berkeley.

Hirata, S. 1983. "Diminutives" and tone sandhi. *Computational Analysis of Asian and African Languages* 21: 43–58. (In Chinese.)

Hong, C. 1982. "Readings in Chinese Linguistics from Different Historical Periods." Nanjing: Jiangsu People's Press. (In Chinese.)

Huang, Y. H., S. Pratoomraj, and R. C. Johnson. 1969. Universal magnitude symbolism. *Journal of Verbal Learning and Verbal Behavior* 8(1): 155–156.

Jakobson, R. 1978. (J. Mepham, trans.) *Sound and Meaning*. London: MIT Press.

Jespersen, O. 1922. *Language: Its Nature, Development and Origin*. London: George Allen and Unwin.

Johnson, R. C., N. S. Suzuki, and W. K. Ohls. 1964. Phonetic symbolism in an artificial language. *The Journal of Abnormal and Social Psychology* 69: 233–236.

Kam, T. H. 1977. Derivation by tone change in Cantonese. *Journal of Chinese Linguistics* 5(2): 186–210.

1980. Semantic-tonal processes in Cantonese, Taishanese, Bobai, and Siamese. *Journal of Chinese Linguistics* 8(2): 205–240.

Klank, L. J. K., Y.-H. Huang, and R. C. Johnson. 1971. Determinants of success in

matching word pairs in tests of phonetic symbolism. *Journal of Verbal Learning and Verbal Behavior* 10: 140–148.

Klatt. D. H. 1973. Discrimination of fundamental frequency contours in synthetic speech: implications for models of perception. *Journal of the Acoustic Society of America* 53: 8–16.

Koriat, A. 1975. Phonetic symbolism and feeling of knowing. *Memory and Cognition* 3(5): 545–548.

Lakoff, Robin. 1975. *Language and a Woman's Place*. New York: Harper & Row.

Lester, D. 1973. Phonetic and graphic symbolism. *Perceptual and Motor Skills* 37: 592.

1974. Symbolism in the Chinese language. *International Journal of Symbology* 5(1): 18–21.

Li, C. N. and Thompson, S. A. 1976. The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* 4: 185–199.

Li, F. K. 1977. *A Handbook of Comparative Tai*. Honolulu: University Press of Hawaii.

Lu, J. 1981. The classification and use of monosyllabic antonyms. In Peking University Chinese Department, *Yuyanxue Luncong* Editorial Committee (eds.) *Collection of Papers in Linguistics*, no. 8. Beijing: Commercial Press, 34–49. (In Chinese.)

Ma, X. 1981. "A Survey of Linguistics." Wuchang: Huazhong Gongxueyuan Press. (In Chinese.)

Maddieson, I. 1978. Universals of tone. In J. H. Greenberg (ed) *Universals of Human Language* II: *Phonology*. Stanford: University Press, 335–365.

Maltzman, I., L. Morrisett, L. and L. Brooks. 1956. An investigation in phonetic symbolism. *The Journal of Abnormal and Social Psychology* 53: 245–251.

Matisoff, J. A. 1973. *The Grammar of Lahu*. University of California Publications in Linguistics 75. Berkeley: University of California Press.

Miron, M. S. 1961. A cross-linguistic investigation of phonetic symbolism. *The Journal of Abnormal and Social Psychology* 62: 623–630.

Morton, E. S. 1977. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *American Naturalist* 111: 855–869.

1986. Sound symbolism and its role in non-human vertebrate communication. Paper given at the Conference on Sound Symbolism, University of California at Berkeley, January 16–18, 1986.

Newman, S. S. 1933. Further experiments in phonetic symbolism. *American Journal of Psychology* 45: 53–75.

Ohala, J. J. 1980. The acoustic origin of the smile. *Journal of the Acoustic Society of America* 68: S33.

1982a. Physiological mechanisms underlying tone and intonation. In H. Fujisaki and E. Garding (eds.), *Preprints of the Working Group on Intonation, Thirteenth International Congress of Linguists, Tokyo*. Tokyo: ICL Editorial Committee, 1–12.

1982b. The phonological end justifies any means. *Preprints of the Plenary Session Papers, Thirteenth International Congress of Linguists, Tokyo*. Tokyo: ICL Editorial Committee, 199–208.

1982c. The frequency code and its effect on certain forms of speech and facial expressions. In A. S. House (ed.) *Proceedings of the Symposium on Acoustics, Phonetics and Speech Modelling*. Paper 23/81. Princeton: Institute for Defense Analysis, 1–31.

1983. Cross-language use of pitch: an ethological view. *Phonetica* 40: 1–18.

1986. Sound symbolism in an ethological context. Paper presented to the Conference on Sound Symbolism, University of California at Berkeley, January 16–18, 1986.

Sapir, E. 1929. A study of phonetic symbolism. *Journal of Experimental Psychology* 12: 225–239.

Saussure, F. de. 1966. *Course in General Linguistics*. Edited by C. Bally and A. Sechehaye in collaboration with Albert Reidlinger. Transl. and notes by Wade Baskin. New York: McGraw-Hill.

Solomon, L. N. 1959. Search for physical correlates to psychological dimensions of sounds. *Journal of the Acoustic Society of America* 31: 492–497.

Taylor, I. K. 1963. Phonetic symbolism reexamined. *Psychological Bulletin* 60: 200–209.

Taylor, I. K. and M. M. Taylor. 1965. Another look at phonetic symbolism. *Psychological Bulletin* 64: 413–427.

1967. Phonetic symbolism in four unrelated languages. *Canadian Journal of Psychology* 16: 344–356.

Tse, J. K-P. 1977. Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language* 5: 191–204.

Tsien-Lee, M. 1969. Sound and meaning in Chinese language: a study of phonetic symbolism. *Psychologica Belgica* 9(1): 47–58.

Vance, T. J. 1976. Tone and intonation in Cantonese. *Phonetica* 33: 368–392.

Wang, W. S-Y. 1979. Language change – a lexical perspective. *Annual Review of Anthropology* 8: 353–371.

Weiss, J. 1963. The role of "meaningfulness" vs. meaning-dimensions in guessing the meanings of foreign words. *The Journal of Abnormal and Social Psychology* 66: 541–546.

1964. Phonetic symbolism re-examined. *Psychological Bulletin* 61: 454–458.

Whitaker, K. P. K. 1955–1956. A study on the modified tones in spoken Cantonese. *Asia Major* 5: 9–36; 184–207.

Wong, M. K. 1982. Tone change in Cantonese. PhD dissertation, University of Illinois at Urbana Champaign.

Woon, W.-L. 1979. A synchronic phonology of Hainan dialect: Part I. *Journal of Chinese Linguistics* 7: 65–100.

Wu, Y. 1984. On register in a tonal theory of Mandarin. Paper presented to the Seventeenth International Conference on Sino-Tibetan Languages and Linguistics, University of Oregon, Eugene, September 7–9, 1984.

Ye, F. and T. Xu. 1981. "An Outline of Linguistics." Beijing: Peking University Press. (In Chinese.)

Yue-Hashimoto, A. O. 1972. *Phonology of Cantonese*. Studies in Yue Dialects 1. Cambridge: University Press.

1980. Word play in language acquisition: a Mandarin case. *Journal of Chinese Linguistics* 8(2): 181–204.R.

Zhong, Q. 1980. *On Chinese Phonetics*. Beijing: Commercial Press.