# 行政院國家科學委員會專題研究計畫　成果報告

## 中文詞彙聲調相對強度之聲學實驗研究

計畫類別：個別型計畫
計畫編號：NSC93-2411-H-040-002-
執行期間：93 年 08 月 01 日至 95 年 01 月 31 日
執行單位：中山醫學大學應用語言學系

計畫主持人：江羅葛
共同主持人：江幸真

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 95 年 5 月 1 日

## ABSTRACT

### Acoustic Experimental Study Of Relative Intensity In Chinese Word Tone

Almost all studies of (Mandarin) Chinese prosody agree that it has lexically specified word tones. The most common view in the current linguistics literature on their form is that pitch is the sole or primary component. A number of studies also see length as being a secondary component. Very few studies suggest that loudness may also be a component.

This study presents original acoustic experimental data on the relative-intensity patterns of tones, with the hypothesis being that loudness is indeed a component. In addition, a review of findings from experimental studies on the duration of tones leads to the conclusion that it is unlikely that length actually is a component of tones, at least in connected speech.

This study's experimental data comes from six sets of four homophonous target words of the same syntactic category (one word for each tone) in a carrier dialog produced by three adult native-speakers of Standard Taiwanese Chinese. The productions were digitally recorded and entered into a computer for analysis of the relative intensity (RI) flows syllable rhymes in the target words. The productions of the target words in each set were compared to see if the tones have systematic word-level RI flows associated with them that could be an acoustic basis for a loudness-pattern component of Chinese word tone.

The results of this study suggest that the four different lexical tones of the target words have distinct RI-based loudness-pattern components.



Keywords: (Mandarin) Chinese, word tone, sentence stress, relative intensity, fundamental frequency, loudness, pitch.

## 中文摘要

### 中文詞彙聲調相對強度之聲學實驗研究

幾乎所有關於中文音韻（prosody）的研究都認為中文具有特定之詞彙聲調。在目前語言學的文獻上，最普遍的論述為：音調（pitch）是唯一或者是最主要的成分。也有很多研究認為音長（length）是次要的成分。另有少數的研究認為響度（loudness）可能是其中的一個成分。

在預設響度確實是中文詞彙聲調的一個成分之假設為前提下，此研究將採集相關資料並驗證中文詞彙聲調相對強度形式之原始聲學實驗資料。除此之外，從此探究中文詞彙聲調音長實驗之研究結果及發現，將可導出結論至少在連續的言談方面，聲調音長(長度)事實上不太可能會是聲調的一個成分。

在資料的設計、採樣、及分析上，本研究的採樣資料是來自三位在台灣講標準中文的台灣人。本資料庫的設計是由 6 組載體形式對話之目標語所組成。各組由四個同音且具有相同句法範疇(一個字一個聲調)之目標語所組成。首先以數位錄音方式儲存、然後輸進電腦軟體作聲學資料分析：分析此研究的目標語音節（節奏音韻）之 relative intensity (RI) flows。每一組目標語所採樣之資料將互相作比較，以整體比較之結果查驗中文在詞彙上聲調是否呈現出有系統的 RI flows 現象。 這即顯示響度為中文詞彙聲調成分之可能。

此研究的結果顯示目標語的四種不同詞彙聲調具有顯著 RI 為基準之響度形式的成分。



關鍵字：中文，詞的聲調，句子重音，相對強度，基本頻率，響度音長，音調

# Acoustic Experimental Study Of Relative Intensity In Chinese Word Tone

## Introduction

This study's experiment is designed to elicit acoustic data bearing on a fundamental issue of prosodic sound in (Mandarin) Chinese: the form of tone.

## Research Question

Specifically, the purpose of the experiment is to examine whether word-level tones have distinctive loudness patterns associated with them, and if so, to determine what they are.

## Literature Review And Rationale

Virtually all studies of Chinese hold that it has four lexically specified tones at the word level, and that most syllables in words are lexically specified for one of these tones (e.g., Garding 1987, Kratochvil 1998, Ramsey 1987, Rumyantsev 1987, Shen 1990 &1993, Whalen & Xu 1992, Yip 1990). Although not always discussed, the function of the tones is generally considered to be to help distinguish or identify morphemes and/or words (e.g., Ramsey, 1987, Rumyantsev 1987, Shen 1990).

Most studies maintain that the form of these tones solely or primarily involves pitch, and hence what might be called "periodicity frequency", not necessarily fundamental frequency – see Moore (1997) on theories of pitch perception (e.g., Rumyantsev 1987, Shen 1990, Whalen & Xu 1992, Yip 1990). Many studies maintain that the form is partially or secondarily based in length (e.g., Rumyantsev 1987, Shen 1990, Whalen & Xu 1992). A few studies hold that loudness is also a component (e.g., Kratochvil 1998, Rumyantsev 1987, Whalen & Xu 1992).

### *Tones And Periodicity Frequency*

The pitch aspect of the form is considered in most studies to consist of pitch patterns and pitch register (e.g., Kratochvil 1998, Shen 1990, Yip 1990). The pitch patterns and pitch register are in turn considered in many studies to consist of dynamic movements – rises, falls, and levels (e.g., Garding 1987, Kratochvil 1998, Ramsey 1987, Rumyantsev 1987, Shen 1990 &1993, Whalen & Xu 1992), and in some studies to consist of static points/levels – that is, of highs and lows (e.g., Yip 1990).

Some F0 (fundamental frequency) flow data supporting that the four lexically specified tones of Chinese have distinctive periodicity-frequency (and presumably hence also distinctive pitch) patterns associated with them can be seen in Coster & Kratochvil (1984: 126), Ho (1976b: 355-356), Howie (1974: 139-147 & 1976: 200-220), Shen (1990: 35), Tseng (1990: 18), and Wong (1953: 274). The data in Coster & Kratochvil (1984) is on words of unspecified numbers of syllables in extemporaneous connected speech. Part of the data in Ho (1976b) is from monosyllabic words in read disconnected speech (Pgs. 355-356, 363); the other data is from syllables in read connected speech with no indication given as to whether the syllables are monosyllabic words or parts of polysyllabic words (Pgs. 356, 363). Judging from Ho (1976a), (a very similar study that is apparently a somewhat less developed earlier version of Ho (1976b) and that is more specific on this point), part of this other data in Ho (1976b) is from monosyllabic nouns in read connected speech

and part of it is on the first syllables of bisyllabic nouns in read connected speech in which the second syllables of those nouns are said by Ho to have had a Tone-4. The data in Howie (1974) is from the first syllables of bisyllabic nouns in read connected speech in which the second syllables of those nouns are said by Howie to have been read with "neutral tones" – i.e., not one of the four lexically specified tones (Pg. 137); the data in Howie (1976) is from monosyllabic words in read disconnected speech. The data in Shen (1990) is on monosyllabic and polysyllabic words in read connected speech. The data in Tseng (1990) is from monosyllabic words or nonsense-words in read disconnected speech, from monosyllabic and/or polysyllabic words in extemporaneous connected speech, or an unspecified mixture of both (Tseng does not make these points clear). The data in Wong (1953) is from monosyllabic words in read disconnected speech.

Not all of these studies have F0 flow data of equal detail, and much of the data is averaged data, which does not show any possible varietal differences among the subjects (such as from different accents or dialects) or variations among the tokens for a tone produced by a subject. However, the F0 data presented in all of the studies is consistent with Tone-1 having a basically level pitch pattern high in the pitch range the speaker is using, Tone-2 having a basically rising pitch pattern (with maybe a slight dip at the beginning) from around mid-low to high in the pitch range the speaker is using, and Tone-4 having a basically falling pitch pattern (with maybe a slight hook at the beginning) from high to around mid-low in the pitch range the speaker is using. For Tone-3, all of these studies present F0 flow data consistent with roughly the first half of Tone-3 having a slightly falling pitch pattern from roughly mid-low to low in the pitch range the speaker is using. For the other part of Tone-3, Coster & Kratochvil (1984) presents F0 data consistent with roughly the second half of Tone-3 having a roughly level pitch pattern low in the pitch range the speaker is using; while the other studies present F0 data consistent with roughly the second half of Tone-3 having a slight rising pitch pattern from low to mid-low in the pitch range the speaker is using.

There are two likely reasons for the difference in the F0 data on the second half of Tone-3s. First, although not necessarily the case, it may be that Tone-3 tends at least sometimes to be said somewhat different in extemporaneous speech than in read speech, and it seems unlikely that many speakers would, or could, read sentences so that they would sound exactly the same as if the sentences had been extemporaneously spoken. The Coster & Kratochvil (1984) data is the only data that is clearly solely on extemporaneous, connected speech, and might be different for that reason. Second, although also not necessarily the case, it may be that not all of the speakers who were the sources of the F0 data of the various studies spoke Chinese in exactly the same way. That is, it may be that not all of these speakers spoke the same accent or even the same dialect of Chinese. It may be that the data-source speaker of Coster & Kratochvil (1984) was speaking a dialect or accent of a dialect of Chinese in which Tone-3s do not have a rising pitch in their second half, at least in extemporaneous, connected speech.

In Coster & Kratochvil (1984), the data-source subject is described as "a native female speaker of BD" (Pg. 120) – where "BD" is short for "Beijing Dialect" (Pg. 119). There are several problems with that description, however. First, no explanation is made of "Beijing Dialect". To someone unfamiliar with (Mandarin) Chinese or various fields related to the study of (Mandarin) Chinese, "Beijing Dialect" in a linguistics paper might reasonably be expected to mean something like the

dialect of (Mandarin) Chinese prevalent in Beijing or the Beijing-area. Unfortunately, however, the terms "dialect" and "language" have often been used by researchers and teachers of Chinese and other related languages, both native speakers and foreigners, not in their linguistics sense but rather in some other, unspecified sense - with, for example, related, but mutually unintelligible languages such as (Mandarin) Chinese and Cantonese described as "dialects" of "Chinese", rather than as different languages of a Sinitic language family. Given this situation, it is not clear what is meant by "Beijing Dialect".

Second, Coster & Kratochvil (1984) does not clarify what is meant by "native speaker". Some Chinese speakers who consider themselves and are considered by others to be native speakers of Chinese learned Chinese in their home as a first language, but many others learned Chinese first at school as a second native language, after first learning a different native language at home. As far as I know, this is true both in China and in Taiwan, where Chinese can be said to be the official language. There may or may not be a difference in the speech of monolingual native speakers of Chinese, bilingual first-language native speakers of Chinese, and bilingual second-language native speakers of Chinese, but details like a subject being bilingual with the language under consideration their first or second native language, should be mentioned in a description of the subject's language status.

Last, no matter what Coster & Kratochvil (1984) means by "native speaker", no details of their data-source subject's background are given to support that she actually reasonably can be taken to be a "native speaker of BD", such as where she grew up. Even assuming that subjects might generally be expected to truthfully report such things, which is not a given, there is no guarantee that a given subject would necessarily be more accurate in their description of this than researchers and teachers of Chinese have been - either as to whether they are bilingual, as to which dialect of Chinese they speak, as to whether they speak Chinese as a native language, or as to whether they speak Chinese as a first language or as a second language.

In Ho (1976b & 1976a), the data-source subjects are merely said to be "five Pekinese speakers" (Pgs. 355 & 4, respectively). It is not clear whether Ho intends "Pekinese" to mean Peking(-area) Chinese, or simply (Mandarin) Chinese - as opposed to Cantonese, for example; and it is not clear how accurate Ho is in that assertion, since no details are given on the subjects' backgrounds.

In Howie (1974), the data-source subject is identified as "a male speaker of Peking Mandarin" who was "born in Peking and completed elementary school there", whose "high school years were spent in Peking and Chungking", and whose "university study was done on Taiwan" (Pgs. 134, 137). Although it not mentioned where the subject was for junior high school (possibly Howie intended "high school years" to cover both junior and senior high school), it seems likely that at or sometime after puberty the subject left Peking and went to Chungking, in Sichwan Province, where the Chinese generally spoken would not necessarily have been the same accent or even dialect as that generally spoken by people in Peking. After that the subject went to Taiwan for college (for four years?), where the Chinese generally spoken would not necessarily have been the same accent or even dialect as that generally spoken by people in Chungking or Peking. After those adolescent years in Chungking and those college years in Taiwan, it does not seem that it would be unusual for the speaker to no longer have spoken the same accent and maybe even to not any longer have spoken

the same dialect as that generally spoken in Peking.

In Howie (1976), the data-source subject is merely said to be "one male speaker of Peking Mandarin" (Pg. 146), with no explanatory or supporting details given for that claim, and with the speaker's native-language status unspecified.

In Shen (19900, the six data-source subjects are described as "six female MC [Mandarin Chinese] informants" (Pg. 14). No mention is made of what dialect of Chinese each spoke and the background information on each was limited to age, place of birth, level of education, "later residence", and "time in U.S.A." - with no explanation given of what "later residence" might mean. From this, of course, it is not at all clear what the Chinese language status of the six subjects is.

In Tseng (1990), the data-source subject is described as "an adult female speaker of the standard Peking dialect" (Pg. 14). From that it is not clear whether the speaker was a native speaker or not, and if a native speaker, whether a monolingual native speaker or a bilingual native speaker, and if a bilingual native speaker, whether a first or second language native speaker. It is also not clear what is meant by "Peking dialect" (the Chinese language, the Peking-area dialect of the Chinese language, or the Peking accent of the Peking-area dialect of the Chinese language?), or what is meant by "standard" in "the standard Peking dialect", especially since judging from the romanization and Chinese-character-derived syllabary used by Tseng in the paper and in some of the test material read by her subject, both Tseng and her subject were very likely at least partially educated in Taiwan, where neither the Peking accent of the Peking-area dialect of Chinese nor the Peking-area dialect of Chinese are common, let alone prevalent.

Although the dialect and native language statuses of the subjects who were the sources of the tones' F0 flow data in the studies are not made clear and apparently are not controlled for in the studies, the tones' F0 flow data reported in the studies is fairly clearly consistent with the view that Chinese lexically specified tones have distinctive pitch patterns associated with them.

*Tones And Duration*

Most studies that view length as being a component of the tones hold that it is the total length of the tones that is significant (e.g., Shen 1990, Whalen & Xu 1992). At least one study - Rumyantsev (1987) - proposes that length is a significant factor in the tones with respect not to total length but instead to inflection points for changes in pitch and loudness, that is, that the relative lengths of the falls and the rises and the places where they occur helps distinguish the tones from each other.

Based on the synthesized results of a number of studies, it cannot be concluded that in non-isolated, connected speech Chinese word tones are even partially or secondarily based in length and/or duration. There seem to be no consistent enough correspondences between Chinese lexically specified tones and their relative lengths or durations such that the tones in connected speech could in general be analyzed as having distinctive durations or lengths. When the four tones all occur on monosyllabic words or morphemes said in isolation, there is some evidence that Tone-3 generally tends to be somewhat longer in duration than the other tones and that Tone-4 less generally but still to some degree tends to be somewhat shorter in duration than the other tones. In connected speech, there is evidence of even less of a connection between duration and the tones.

Ho (1976b: 356, 363), Tseng (1990: 22-23), and Whalen and Xu (1992: 27, 29) all present data in which different lexical tones said in isolation tend to have different durations associated with them – non-duration details on the data from these studies are as described earlier, in the section on F0 data. Except for Tone-3 generally having the longest duration of the four lexical tones, however, the results of those studies do not combine into a very consistent duration order for the four lexical tones in non-connected speech. Part of the results from Ho (1976b) has the tones' duration order from longest to shortest as Tone-3, Tone-2, Tone-1, and Tone-4 (Pg. 356), and the other part has it as Tone-3, Tone-1, Tone-2, and Tone-4 (Pg. 363). Tseng (1990) has fairly but not completely similar results, with Tone-3 as the longest, Tone-1 and Tone-2 varying as to which was second and third in duration, and Tone-4 as the shortest (Pg. 22), but with a fair amount of overlap in the duration ranges of Tone-1 and Tone-4 for part of that data (Pg. 23). Whalen & Xu (1992) has somewhat similar but somewhat dissimilar results: part of their results has the tones' duration order from longest to shortest as Tone-3, Tone-2, Tone-1, and Tone-4 (Pg. 27) - similar to the Ho (1976b) and Tseng (1990) results; but part of their results has the order as Tone-3, Tone-2, and Tone-4/Tone-1 (Pg. 29). Putting these results together, all that can be said is that when the tones all occur on monosyllabic words or morphemes said in isolation, Tone-3 generally tends to be somewhat longer in duration than the others and Tone-4 less generally but still to some degree tends to be somewhat shorter in duration than the others.

In connected speech, the four tones do not seem to have even that much duration order. Ho (1976b) gives averaged results in which all four tones are basically the same duration in unspecified read connected speech environments (Pgs. 356, 363). Coster & Kratochvil (1984) gives averaged results from extemporaneous, connected speech in which the tones' duration order from greatest to least is Tone-2, Tone-1, Tone-3, and Tone-4 - with Tone-3 thus third in duration, rather than first as in the unconnected speech studies' findings (Pg. 126). Shen (1990) gives averaged results from read connected speech in which the tones' duration order is Tone-3 (the longest) and then Tone-1/Tone-2/Tone-4 (all tied for second, third, and fourth) in "sentence-final position in statements ... and [polar] questions" (Pg. 35). Tseng (1990) gives averaged results from extemporaneous, connected speech in which the tones' duration order from longest to shortest is Tone-3, Tone-4, and Tone-2 (no data is given for Tone-1), but with a great deal of overlap in the duration ranges of all three of those tones (Pg. 23).

As noted earlier, none of these studies is very clear as to what dialect of Chinese each of their data-source subjects spoke or their native-language(s) background or status, and apparently did not control for those factors. This means that not all of the various results of these studies are necessarily on the same dialect of Chinese, or from similar-background native speakers. Even so, the fact that there is a fair amount of overlap in their duration ranges and that there appear to be no durational distinctions in at least some connected speech environments makes it seem unlikely that Chinese lexical tones have a general set duration order – although it may be that in certain environments and/or conditions certain tones tend to (but apparently need not) have relative durations (and hence lengths) associated with them. The tones' duration data reported in the studies, then, seems fairly clearly consistent with the view that Chinese lexical tones are not morphologically specified for length, but that their length depends on factors not related to their morphological specifications.

The proposal in Rumyantsev (1987) that length is a significant factor in tones with respect to inflection points for changes in pitch and loudness, rather than with respect to overall length, seems more likely. As noted, this idea basically is that two tones might for example both be able to be said with an initial fall and then a later rise in, say, pitch, and still be differentiable on the basis of obligatory differences in the relative lengths and starts and ends of the falls and the rises, that is, on the basis of differences in how long they are in the tones and where they occur in the tones.

*Tones And Relative Intensity*

As noted, the most common view in the literature is that Chinese lexically specified tones have pitch patterns but no loudness patterns. Although not very definitive, some evidence that they have RI-flow (and hence loudness) patterns as well, however, can be found in Coster & Kratochvil (1984: 126), Ho (1976b: 357, 360-361), Howie (1976), and Whalen & Xu (1992: 27, 29). These studies all report finding distinctive "amplitude" flows associated with different tones. Ho (1976b) and Coster & Kratochvil (1984) do not specify exactly what they mean by "amplitude", but Whalen & Xu (1992) describes "amplitude" as being measured in decibels, meaning that it presumably actually is not amplitude (neither sound wave amplitude nor sound wave pressure amplitude), but rather sound relative-intensity (RI) level (non-relative-intensity details on the data in these studies are as described earlier, in the section on F0 data).

Howie (1976) reports on two perceptual experiments on identifying the four lexical tones from "citation" monosyllable words/morphemes, presumably meaning that the tones are from read disconnected speech. In one experiment the F0 of the tones was replaced by a monotone, and in the other the F0 flow was replaced by an apparently constant "hiss" - with "amplitudes" and durations apparently unchanged. In the first experiment, the overall correct tone identification is 30.8% and in the second it is 39.2%, although the percent-correct breakdown varies considerably from experiment to experiment and from tone to tone (Pgs. 232-236).

Whalen & Xu (1992) also reports on two perceptual experiments on identifying the four lexical tones from monosyllable words/morphemes in read disconnected speech, in both of which the F0 flows of the tones and the vowel spectrums were replaced by "signal-correlated noise" - the analog-recorded speech signal was first digitized and then the signs of half of the samples selected at random were reversed, leaving an unspecified syllable "amplitude" contour unchanged for both experiments and leaving syllable duration unchanged for the first experiment but changing syllable duration to be the same for all syllables in the second experiment (Pgs. 26, 29). The percent-correct identifications for Tone-1s, Tone-2s, Tone-3s, and Tone-4s were fairly high in both experiments: in the first respectively at 38.5%, around 80%, around 95%, and around 85%, respectively (Pg. 28); and in the second at 45%, 55.3%, 69.5%, and 92.3%, respectively (Pgs. 30-31). To explain the difference between these results and those of Howie (1976) and certain other studies, Whalen and Xu (1992) suggests that "when an Fo pattern is present, it is difficult to ignore it in making a tone judgement ... [but when] there is no Fo ... the amplitude contour can be effective in tone identification" (Pg. 28).

Whalen and Xu (1992), like Ho (1976b), Howie (1976), and Coster and Kratochvil (1984), does not specify or establish the claimed dialect and native-speaker status of the subjects very

clearly. Without the relevant details on the subjects' backgrounds, it is not clear exactly what their accent, dialect, native-speaker status, and first and second language status were, or if they spoke only one language natively or perhaps more. Still, as with the F0-flow data, even with these unclarities on the subjects' Chinese-speaker statuses, and even with the unclarity of exactly what is meant by "amplitude" in Ho (1976b) and Howie (1976), the "amplitude" and apparently RI-flow data of these studies seems fairly clearly consistent with the view that Chinese lexical tones not only have distinctive pitch-flow patterns, but also have distinctive loudness-flow patterns - although less salient than the pitch-flow patterns associated with them.

**Method**

Two adult native-speakers of Standard Taiwanese Chinese were asked to individually read over and become familiar with the cue cards for six four-word sets of target words and the script of a mini-dialog of carrier utterances, all written in Chinese. Then with the help of cue cards and the dialog script, they were asked to say the dialog for each target word twice, with a shuffled order of presentation of the target word cue cards. They were asked to say the utterances in the dialog at what they consider a normal rate of speaking and in a non-ironic, non-emphatic, and more-or-less neutrally interested way. When they forgot the dialog script, a brief break was taken and the script was reviewed before restarting. The productions were digitally recorded using a 20-20000 hertz headset microphone, and entered into a computer for analysis of the RI flows of the rhymes of the syllables in the target words (see Howie 1974 on Chinese word-tones and rhymes). Praat speech-analysis software was used to obtain sound waveforms, spectrograms, RI measurements, and f0 measurements for the tokens (see Boersma 1993 on f0 extraction in Praat).

Each target-word set consisted of four words of the same syntactic category, which were syllabically and phonemically identical but different in lexically specified tone. The experimental dialog was designed so as to be able to compare sets of target words with the four lexically specified tones (Tones 1-4) in utterance positions that receive sentence stress and that are controlled for intonation, to see if the tones have systematic word-level RI flows associated with them that could be an acoustic basis for a loudness-pattern component of word tone.

**Results & Discussion**

As illustrated in the following graphs of the relative intensity (db) flows of a typical four-word set, the results support the conclusion that the four lexically specified tones in Chinese have distinct RI-based loudness pattern components.

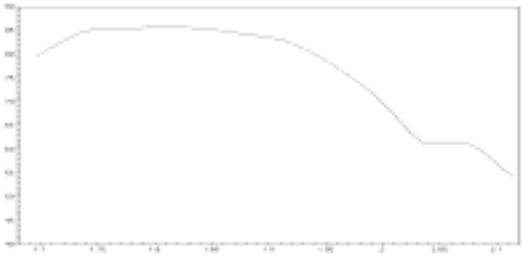
Before comparing the eight graphs' RI flows for tone-based differences, it should be noticed that the differences in intonation seem to mainly affect the second halves of the tones.

Tones 2 and 3 both have two dips in their intensity, regardless of which intonation is used - although in Tone 2 the dips are spaced farther apart and in Tone 3 are spaced closer together. Tones 1 and 4 do not really have dips with declarative intonation and each has one dip with polar interrogative intonation – although the dip is near the end for Tone 1 and about in the middle for Tone 4.

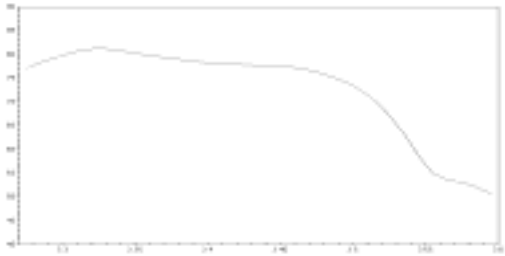
Tone 1 has a slight fall over more than half of its course regardless of the intonation, while

Tone 2 is almost level over more than half of its duration regardless of the intonation. If listeners allow for a small downdrifting of intensity over the course of a tone, the slight fall of Tone 1 could be interpreted as level and the levelness of Tone 2 as a rise. Tone 3 has a slight fall during its first half for both intonations, which as noted might be interpreted as being more or less level by listeners. The second half of Tone 3 maintains the slight fall with polar interrogative intonation but falls sharply with declarative intonation, Tone 4 has a relatively steady fall over most of its course regardless of intonation, although the rate of fall in its first half is slower with polar interrogative intonation than with declarative intonation.

**Illustrative Intensity(db)-Over-Time Graphs: A Four-Word Set Said With Two Intonations**

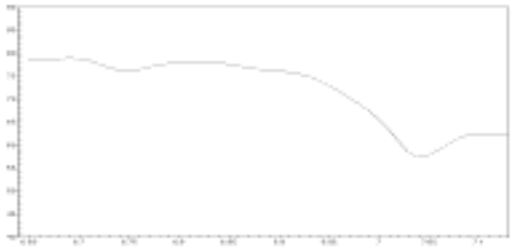


ju1 ("pigs") – polar Q intonation

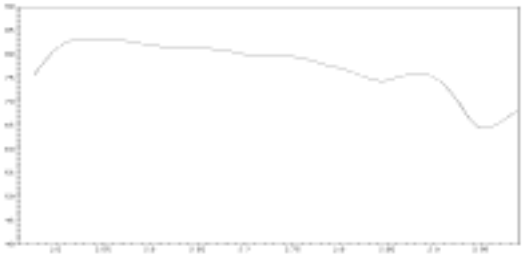


ju1 ("pigs") – statement intonation

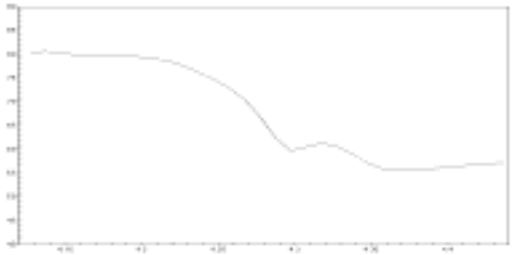
ju2 ("bamboo") – polar Q intonation



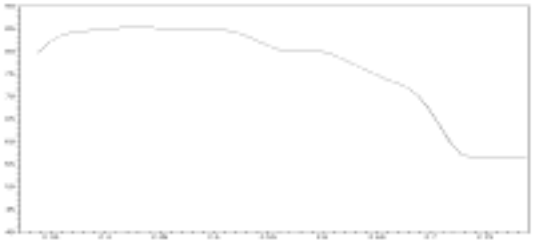
ju2 ("bamboo") – statement intonation



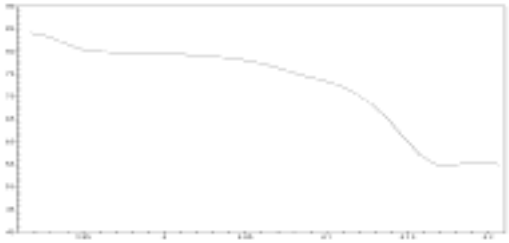
ju3 ("dusters") – polar Q intonation



ju3 ("dusters") – statement intonation



ju4 ("chopsticks") – polar Q intonation



ju4 ("chopsticks") – statement intonation

## References


Boersma, Paul (1993). "Accurate Short-Term Analysis Of The Fundamental Frequency And The Harmonics-To-Noise Ratio Of A Sampled Sound". IFA Proceedings: 17: 97-110. Amsterdam, The Netherlands: Institute Of Phonetic Sciences, University Of Amsterdam.

Coster, D.C. & Kratochvil, Paul (1984). "Tone And Stress Discrimination In Normal Beijing Dialect Speech". Beverly Hong (Ed.) New Papers On Chinese Language Use (Contemporary China Papers No. 18). Contemporary China Centre Research School Of Pacific Studies, The Australian National University, Canberra, Australia.

Garding, Eva (1987). "Speech Act And Tonal Pattern In Standard Chinese: Constancy And Variation". Phonetica 44 13-29.

Ho, Aichen Ting (1976a). "Mandarin Tones In Relation To Sentence Intonation And Grammatical Structure". The Journal Of Chinese Linguistics 4:1:.

Ho, Aichen Ting (1976b). "The Acoustic Variation Of Mandarin Tones". Phonetica 33: 353-367.

Ho, Aichen Ting (1977). "Intonation Variation In A Mandarin Sentence For Three Expressions: Interrogative, Exclamatory, and Declarative". Phonetica 34: 446-457.

Howie, John Marshall (1976). Acoustical Studies Of Mandarin Vowels And Tones. Cambridge, U.K.: Cambridge University Press.

Howie, John Marshall (1974). "On The Domain Of Tone In Mandarin: Some Acoustic Evidence". Phonetica 30: .

Kratochvil, Paul (1998). "Intonation In Beijing Chinese". Hirst, Daniel & Di Cristo, Albert (Eds.). Intonation Systems: A Survey Of Twenty Languages: 417-431. Cambridge, U.K.: Cambridge University Press.

Moore, Brian C. J. (1997). An Introduction To The Psychology Of Hearing. San Diego, California, U.S.A.: Academic Press, Incorporated.

Ramsey, S. Robert (1987). The Languages Of China. New Jersey, U.S.A.: Princeton University Press.

Rumyantsev, M. K. (1987). "Chinese Tones And Their Duration". Proceedings Of The 11th International Congress Of Phonetic Sciences: 1: 166-169.

Shen, Xiao-nan Susan (1993). "Relative Duration As A Perceptual Cue To Stress In Mandarin". Language And Speech: 36: 4: 415-433.

Shen, Xiao-nan Susan (1990). The Prosody Of Mandarin Chinese (University Of California Publications In Linguistics; V.118). Berkeley, California, U.S.A.: University Of California Press.

Tseng, Chiu-yu (1990). An Acoustic Phonetic Study On Tones In Mandarin Chinese (Academia Sinica Institute Of History And Philology Special Publications No.94). Taipei, Taiwan R.O.C.: Academia Sinica.

Whalen, D. H. & Xu, Xi (1992). "Information For Mandarin Tones In The Amplitude Contour And In Brief Segments". Phonetica 49: 25-47.

Wong, Helen (1953). "Outline Of The Mandarin Phonemic System". Word 9: 3: .

Yip, Moira Jean Winsland (1990). The Tonal Phonology Of Chinese. New York, NY, U.S.A.: Garland Publishing, Inc..

**Project Evaluation**

The realized project was basically the same as the proposed project, with only a few minor adjustments of scale, and after some additional reduction of the data to significant numerical measures and statistical treatment, is quite publishable. The research carried out in this project has value with respect to the acoustic basis of Chinese word tone and prosody and with respect to the acoustic basis of linguistic prosody in general. In addition, the research findings can be applied to the theoretical analysis of Chinese word tone and other aspects of Chinese prosody, and can also be further applied to the theoretical analysis of other languages with different prosodic systems, such as English and Japanese, and thus to an analysis of human language prosody in general.