

Compiling Word Lists with Japanese Word Accent and Chinese Tone for Testing and Training of Discrimination Ability of Change in Voice Pitch Through Cochlear Implant

Liming Yang[†] and Shizuo Hiki[‡]

[†] Institute of Language Teaching, Waseda University, Nishiwaseda, Shinjuku-ku, Tokyo, Japan

[‡] School of Human Sciences, Waseda University, Mikajima, Tokorozawa, Saitama, Japan

E-mail: [†] youl@waseda.jp, [‡] hiki@waseda.jp

Abstract

A system has been designed for testing accurately and training effectively the discrimination ability of change in voice pitch through cochlear implant. The system utilizes speech samples in which change in voice pitch is controlled precisely by a computer program of speech analysis-synthesis, and cooperates with computer-assisted instruction in order to minimize the steps before achieving a goal. Japanese word accent of Tokyo dialect and syllabic tone of Standard Chinese were taken as the subjects for discrimination of change in voice pitch. Compilation of the word lists of the speech samples was based on the following phonetic considerations: 1) phonemic structure in the syllable 2) balance among phonemes in terms of their characteristics of both acoustical analysis and perceptual test, and 3) occurrence frequency of phonemes and syllables.

1 Purpose of the Study

1.1 Cochlear implant device

In the cochlear implant, speech sound is converted into an electrical signal and transmitted to the electrodes implanted in the inner ear, in order to stimulate directly the auditory nerve of a deaf person. In the device, speech sound through microphone is input to a band-pass filter bank. The rectified and smoothed output is fed to the speech signal processor, where the spectral features are modified to make the information of formants and voice pitch easier to be perceived. Then, the signal is encoded to map to the electrodes, transmitted transcutaneously in radio frequency, decoded and fed to the electrodes.

Various type of cochlear implant devices have been developed in the United States, Austria and European countries, and clinically applied effectively to more than ten thousand deaf persons around the world.

1.2 Role of voice pitch

In the recent years, the clinical application is rapidly spreading to Asian countries such as Korea, China, Taiwan and Japan as well. Information conveyed by the voice pitch of the speech sound of the languages spoken in these countries, however, is quite different from that of English and European languages.

The segmental features conveyed by the temporal pattern of voice pitch, in addition to the supra-segmental features in sentence intonation and phrase emphasis, play an important role in speech recognition, especially when combined with the phonemic information obtained by spectral structures and speech-reading (lip-reading). As the transmission quality of the voice pitch of speech sound depends on setting of the processor of each cochlear implant device, it is necessary to measure precisely the discrimination ability for change in voice pitch by cochlear implantees, and to make maximum use of it.

For this purpose, a system has been designed for testing accurately and training effectively discrimination ability of change in voice pitch through cochlear implant.

2 Change in Voice Pitch

2.1 Required resolution

By referring to the data of acoustical analysis of the temporal pattern of voice pitch in Japanese and Chinese speech sounds and their perceptual responses by the normal hearing listeners [1, 2, 3], the required resolution for the voice pitch can be summarized as follows:

- 1) The voice pitch difference between low and high contrast in syllabic segment associated with the Japanese word accent is two whole steps in diatonic scale, with a perceptual resolution of 1/2 of whole step.
- 2) The voice pitch transition in a syllabic segment associated with the Chinese tones is two whole step rise / 0.4 second, and three whole step fall / 0.2 second.
- 3) The typical duration of the syllabic segment is 0.4 seconds, with a perceptual resolution of 0.1 second.

2.2 Control parameters

Parameters to be controlled for the temporal pattern of voice pitch are; a) average voice pitch (high/normal/low) for each of female and male, b) range of voice pitch (wide/normal), c) contrast among temporal patterns of voice pitch (normal/reduced), and d) tempo (normal/fast).

3 CAI Program

3.1 Speech analysis-synthesis

The system utilizes a computer program of speech analysis-synthesis, STRAIGHT developed in ATR

(Advanced Telecommunications Research Institute International) Human Information Processing Research Laboratories by H. Kawahara of Faculty of Engineering, Wakayama University, Japan. In this program, such acoustical parameters as fundamental frequency, intensity and duration of voice source can be controlled precisely.

3.2 Strategy of testing and training

The system cooperates with computer-assisted instruction. In this program, the speech samples are presented either in single or pair, and either once or repeatedly. Response of the discrimination by means of identification or comparison for the stimulus can be reported either orally or through a touch screen. Picture of the stimulus word can be displayed on the screen, especially for young children.

The voice pitch patterns of the synthetic speech are modified according to the history of the response. The degree of the difficulty of discrimination of the next stimulus stays the same or becomes less after an incorrect response, but becomes more after a correct response. The order of presentation of the combination of the control parameters is selected according to the history of response of the listener, so that the number of steps of presentation becomes minimum before finding a reliable test result, or achieving a training goal.

The visual display of the temporal pattern of voice pitch on the computer screen can be utilized in the training procedure for enhancing the auditory stimulus.

4 Japanese Word Accent

4.1 Phonetic considerations

Japanese word accent and Chinese syllabic tone were taken as the most important subjects for discrimination of the change in voice pitch. Compilation of the word lists of speech samples to be installed in this testing and training system was based on the following phonetic considerations; 1) phonemic structure in syllable in each language, 2) balance among phonemes in terms of their characteristics of both acoustical analysis and perceptual test, and 3) occurrence frequency of phonemes and syllables.

4.2 Contrast of word accent

As for the Japanese word accent, groups of bi-syllabic words which have the same phoneme series but different meanings corresponding to the word accent were selected. The possible contrast is whether the down skip of word accent from high to low voice pitch is after the first syllable, the second syllable, or none. Among the bi-syllabic words found in *Japanese Word Accent Dictionary* (edited by Japan Broadcasting Corporation, 1967, in Tokyo dialect), number of groups with the contrast of after the first syllable versus none was most, but after the first versus second syllables reduced by half, and after the second syllable versus none reduced by one-fourth.

Examples of the groups of bi-syllabic words with those contrast of Japanese word accent are shown in Table 1. The words with unvoiced consonants, in which the change in voice pitch is more difficult to perceive than voiced ones, are introduced at the second stage of testing and training procedure.

5 Chinese Tone

5.1 Combination of tones in bi-syllabic words

As for the syllabic tone in Standard Chinese, groups of mono-syllabic words which have the same phoneme series but different meanings corresponding to the tones, and groups of bi-syllabic words which have the same phoneme series either in the first or second syllables, but different meanings corresponding to the combination of tones, were selected.

In the bi-syllabic words, there are fifteen combinations of four kinds of tones, namely, Tone-1 (high flat), Tone-2 (rising), Tone-3 (low and rising) and Tone-4 (falling), in the first and second syllables. (One combination is missing because of the tone change from Tone-3+Tone-3 to Tone-2+Tone-3). The tones are weakened in the second syllable in case of the postpositional particles type utterances (light tone, denoted Tone-5 hereafter).

Examples of the groups of mono-syllabic and bi-syllabic words are shown in Table 2. Most of the syllables consisted of either vowel or voiced consonant + vowel.

5.2 Use of statistical data of tones

Results of perceptual test on the Chinese tone by both native speakers and foreign learners, and the table of syllables constructed by Chinese phonogram [pinyin] in reference to Japanese kana representation, proposed by the present authors, were utilized in the compilation of the word lists.

Occurrence probability of combination of tones was also utilized systematically, by analyzing the statistical data found in *The Grammatical Knowledge-base of Contemporary Chinese* (ed. S-W. Yu, Tsinghua University Press, China, 1998), in choosing words suitable for daily use and especially for children

References

- [1] C-K. Chuang, S. Hiki, T. Sone and T. Nimura, "Acoustical features of the four tones in monosyllabic utterances of Standard Chinese," *J. Acoust. Soc. Japan*, Vol. 31, pp. 369-380, 1975 (in Japanese).
- [2] L-M. Yang, "Experimental research into the perception of Chinese tones," *Bulletine, Res. Inst. Human Sciences, Meiji University*, pp. 293-307, 1999 (in Japanese).
- [3] S. Hiki, K. Imaizumi and Y. Fukuda, "Design of cochlear implant device for transmitting voice pitch information in speech sound of Asian languages," *Proc. 5th Int'l Conf. Spoken Language Processing*, #1091, Sydney, Australia, 1998.

Table 1. Examples of groups of bi-syllabic words with contrast of Japanese word accent (in Tokyo dialect).

[↘]: Downskip of word accent from high to low voice pitch

| after 1st syllable | after 2nd syllable | none |
|--------------------|---------------------|----------------------------|
| [a↘ki] | [a ki↘] | [a ki] |
| 秋 autumn | 飽き tiresomeness | 空き an opening |
| [o:↘me] | [o: me↘] | [o: me] |
| 青梅 name of a town | 多め a larger portion | 大目 tolerate |
| [ka↘ki] | [ka ki↘] | [ka ki] |
| 牡蠣 an oyster | 垣 a fence | 柿 a persimon |
| [tsu↘ru] | [tsu ru↘] | [tsu ru] |
| 鶴 a crain | 弦 a string | 釣る fish |
| [ha↘fi] | [ha fi↘] | [ha fi] |
| 箸 chopticks | 橋 bridge | 端 edge |
| [ma↘ku] | [ma ku↘] | [ma ku] |
| 撒く sow seeds | 幕 a curtain | 巻く wind up |
| [rjo:↘ba] | [rjo: ba↘] | [rjo: ba] |
| 良馬 a good horse | 獵場 a hunting ground | 両刃 a double edged sword |

| after 1st syllable | after 2nd syllable |
|------------------------|----------------------|
| [a↘sa] | [a sa↘] |
| 朝 morning | 麻 hemp |
| [a↘fi] | [a fi↘] |
| 葦 a reed | 足 a foot |
| [i↘fi] | [i fi↘] |
| 医師 a medical doctor | 石 a stone |
| [i↘fi] | [i fi↘] |
| 位置 a position | 一 one |
| [i↘ma] | [i ma↘] |
| 今 now | 居間 living room |
| [ka↘fi] | [ka fi↘] |
| 価値 value | 勝ち a win |
| [ka↘me] | [ka me↘] |
| 亀 a tortoise | 瓶 an earthenware pot |
| [ka↘ta] | [ka ta↘] |
| 肩 the shoulder | 型 a model |
| [ka↘mi] | [ka mi↘] |
| 神 God | 髪 hair |
| [fi↘su:] | [fi su:↘] |
| 始終 always | 四十 forty |

| after 2nd syllable | none |
|--------------------|---------------------|
| [ai fo:↘] | [ai fo:] |
| 相性 affinity | 愛称 a nickname |
| [i ppo:↘] | [i ppo:] |
| 一方 the other side | 一報 the first report |
| [u e↘] | [u e] |
| 飢え hunger | 上 a higher place |
| [fo tai↘] | [fo tai] |
| 所帯 a household | 書体 a type face |
| [ɕo: dan↘] | [ɕo: dan] |
| 冗談 a joke | 上段 the upper row |

| after 1st syllable | none |
|---------------------------|---------------------------------|
| [a↘kka] | [a kka] |
| 悪貨 bad money | 悪化 deterioration |
| [a↘me] | [a me] |
| 雨 rain | 飴 a candy |
| [i↘ken] | [i ken] |
| 意見 an opinion | 違憲 unconstitutional |
| [i↘ppai] | [i ppai] |
| 一杯 a cup | 一敗 one defeat |
| [i↘ru] | [i ru] |
| 炒る roast | 居る stay |
| [ka↘u] | [ka u] |
| 飼う raise an animal | 買う purchase |
| [kai↘ho:] | [kai ho:] |
| 介抱 nursing | 開放 open |
| [ka↘ji] | [ka ji] |
| 菓子 confectionary | 貸し a debt |
| [ka↘kko:] | [ka kko:] |
| 郭公 a cockoo | 格好 a figure |
| [ka↘ma] | [ka ma] |
| 鎌 a sickle | 釜 a kettle |
| [ka↘ri] | [ka ri] |
| 雁 a wild goose | 仮 temporary |
| [kan↘ji] | [kan ji] |
| 幹事 an organizer | 漢字 Chinese charactors |
| [kan↘ɰo] | [kan ɰo] |
| 看護 the care | 漢語 a Chinese expression |
| [kan↘ɕo:] | [kan ɕo:] |
| 勘定 counting | 感情 emotion |
| [ɕi↘fin] | [ɕi fin] |
| 自身 oneself | 地震 earthquake |
| [fi↘se:] | [fi se:] |
| 四声 the four tones | 姿勢 a posture |
| [ɕi↘dai] | [ɕi dai] |
| 次代 the next generation | 時代 an epoch |
| [ɕi↘do:] | [ɕi do:] |
| 児童 a child | 自動 automatic |
| [fo:↘ɰai] | [fo: ɰai] |
| 生涯 a lifetime | 障害 a barrier |
| [ɕo↘fi] | [ɕo fi] |
| 女子 girls | 助詞 a postpositional particle |
| [ɕo:↘bu] | [ɕo: bu] |
| 上部 the upper part | 丈夫 healthy |
| [fo:↘nin] | [fo: nin] |
| 商人 a marchant | 証人 a witness |
| [fi↘ro] | [fi ro] |
| 白 white | 城 a castle |
| [tsu↘ju] | [tsu ju] |
| 露 dew | 梅雨 the rainy season |
| [ne↘ru] | [ne ru] |
| 練る knead | 寝る lie down |
| [hon↘to:] | [hon to:] |
| 本島 the main island | 本当 truth |

Table 2. Examples of groups of mono-syllabic and bi-syllabic words with contrast of tones (in Standard Chinese).

A. Groups of mono-syllabic words having the same phoneme series but different meanings corresponding to the tones.

| | Tone-1 | Tone-2 | Tone-3 | Tone-4 |
|----------------------------------|------------------|-------------------|------------------|---------------------------|
| A-1. Syllable consisted of vowel | | | | |
| 1) | [wā] 蛙 frog | [wá] 娃 baby | [wǎ] 瓦 tile | [wà] 袜 socks |
| 2) | [yī] 一 one | [yí] 姨 aunt | [yǐ] 椅 chair | [yì] 亿 hundred million |
| 3) | [wū] 屋 house | [wú] 无 nothing | [wǔ] 五 five | [wù] 物 matter |
| 4) | [yīn] 音 sound | [yín] 银 silver | [yǐn] 饮 drink | [yìn] 印 stamp |

A-2. Syllable consisted of voiced consonant + vowel

| | | | | |
|----|---------------------|----------------------|--------------------|--------------------|
| 1) | [lēi] 勒 strangle | [léi] 雷 thunder | [lěi] 垒 rampart | [lèi] 泪 tear |
| 2) | [liū] 溜 slide | [liú] 留 remain | [liǔ] 柳 willow | [liù] 六 six |
| 3) | [miāo] 喵 mew | [miáo] 苗 seedling | [miǎo] 秒 second | [miào] 庙 temple |
| 4) | [māo] 猫 cat | [máo] 毛 hair | [mǎo] 铆 rivet | [mào] 帽 cap |

B. Groups of bi-syllabic words with the syllables consisted of voiced consonant + vowel

B-1. Contrast of the four tones in the second syllables. Phoneme series of the first syllables are the same.

| 1st syllable | 2nd syllable | | | |
|--------------|--------------------------|--------------------------|---------------------------|--------------------------------|
| | Tone-1 | Tone-2 | Tone-3 | Tone-4 |
| Tone-1 | [māowō] 猫窝 cat bed | [māomáo] 猫毛 cat hair | [māoyǎn] 猫眼 cat eye | [māoròu] 猫肉 cat meat |
| Tone-2 | [yúyīng] 鱼鹰 cormorant | [yúyuán] 鱼圆 fish ball | [yúwěi] 鱼尾 tail | [yúròu] 鱼肉 fish |
| Tone-3 | [yèmāo] 野猫 stray cat | [yèrén] 野人 buffalo | [yèmǎ → yémǎ] 野马 horse | [yèwèi] 野味 game |
| Tone-4 | [ròusōng] 肉松 starch | [ròuwán] 肉丸 meat ball | [ròuyǎn] 肉眼 naked eye | [ròumiàn] 肉面 meat buckwheat |

B-2. Contrast of the four tones in the first syllables. Phoneme series in the second syllables are the same.

| 1st syllable | 2nd syllable | | | |
|--------------|---|----------------------|-----------------------|--------------------|
| | Tone-5: Tones are weakened because of the postpositional particles type utterances. | | | |
| Tone-1 | [wūzi] 屋子 room | [lánzi] 篮子 basket | [lǐngzi] 领子 collar | [wàzi] 袜子 socks |
| Tone-2 | [mōya] 摸呀 touch | [máiya] 埋呀 bury | [mǎiya] 买呀 buy | [màiya] 卖呀 sell |
| Tone-3 | [mōma] 摸吗 touch | [máima] 埋吗 bury | [mǎima] 买吗 buy | [màima] 卖吗 sell |
| Tone-4 | [mōle] 摸了 touched | [máile] 埋了 buried | [mǎile] 买了 bought | [màile] 卖了 sold |

C. Groups of bi-syllabic words with combination of the four tones in the first and second syllables.

| 1st syllable | 2nd syllable | | | | |
|--------------|-------------------------------|---------------------------------|----------------------------------|------------------------|---------------------|
| | Tone-1 | Tone-2 | Tone-3 | Tone-4 | Tone-5 |
| Tone-1 | [sāntān] 三天 for three days | [sānnián] 三年 for three years | [sānwǎn] 三碗 three bowls | [sānyuè] 三月 March | [sānge] 三个 three |
| Tone-2 | [shítān] 十天 for ten days | [shínián] 十年 for ten years | [shíwǎn] 十碗 ten bowls | [shíyuè] 十月 October | [shíge] 十个 ten |
| Tone-3 | [wūtiān] 五天 for five days | [wūnián] 五年 for five years | [wūwǎn → wúwǎn] 五碗 five bowls | [wūyuè] 五月 May | [wūge] 五个 five |
| Tone-4 | [liùtiān] 六天 for six days | [liùnián] 六年 for six years | [liùwǎn] 六碗 six bowls | [liùyuè] 六月 June | [liùge] 六个 six |