

日本語母語話者による中国語声調の知覚と生成

楊 姝怡 (神戸大学大学院)・山田 玲子 (ATR/神戸大学)
syuiyang@gmail.com

1. はじめに

中国語には四つの声調の区別があるが(Chao, 1948), 非母語話者は知覚と生成の両面で声調を混同する場合がある(Chuang *et al.*, 1972; Kiriloff, 1969; Wang *et al.*, 1999; Wang *et al.*, 2003; Yang & Akahane-Yamada, 2016; 他). また, その混同の割合は声調によって異なる. Wang *et al.* (1999) はアメリカ英語母語話者を対象として, 四つの声調のうちの2つを組み合わせたミニマル・ペアの単音節語(例えば, bēi, bèi)を刺激語とした二肢強制選択課題の知覚実験を行った. その結果, 正答率は一声, 三声>四声>二声の順に低下した. また, Wang *et al.* (2003)ではアメリカ英語母語話者に単音節語を発音させ, その音声を中国語母語話者に一声~四声, または「どれでもない」の5つの選択肢から判定させた. その結果, 発話の発音明瞭度(話者が発音した声調と判定者が選んだ声調が一致した率)は一声>四声>二声>三声の順に低下した.

本研究では日本語母語話者を対象として, 中国語声調の知覚実験と発音の収録を行い, 声調による難易度およびその知覚と生成の関係について検討した.

2. 方法

2.1. 実験参加者

日本語母語話者 30 人(男性 14 人, 女性 16 人; 18~28 歳, 平均 21 歳)が参加した. 全員中国語学習経験はなかった. また, 中国語を使う地域での一か月以上の居住経験もなかった.

2.2. 刺激音

4 つの声調で対立する中国語単音節語(例えば, bā, bá, bǎ, bà)を 1 組とし, 36 組 144 語を刺激語とした. 刺激語を 1 人の中国語母語話者(女性, 23)が読み上げ, サンプリング周波数 44.1kHz, 量子化精度 16bit でデジタル化し, 音声ファイルとして格納したものを音声刺激とした.

2.3. 手続き

実験開始時に, 中国語-日本語のバイリンガルである実験者が中国語声調の説明が日本語で記されたプリントを見せ, 説明を読み上げるとともに 4 つの声調の音声を聞かせつつ, 声調の違いを教示した. この教示の後, 実験を開始したが, 実験では刺激の提示, 反応の取得はすべてパソコンで制御した.

知覚実験では, 実験参加者は静かな部屋でパソコンに向かってヘッドホンを付けて座り, キーボードを使って回答した. 各トライアルでは 1 つの刺激語に対してそれと声調で対立する 3 語をディストラクタとし, 4 つの選択肢がピンインでパソコンのディスプレイ上に表示された(図 1). 選択肢はトライアル毎にランダムな順に配置された. 実験参加者には聞

こえた音声と一致した選択肢を選ぶように教示した。刺激音はトライアルの冒頭で 1 回だけ再生され、繰り返し聞くことはできなかった。144 の刺激音をランダムな順で呈示した。

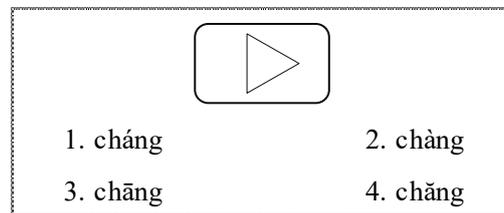


図 1: 知覚実験の画面例

音声の収録においては実験参加者は刺激音を聴取し、それをまねして発音した。実験参加者には声調の変化に注意して発音するように、あらかじめ教示した。その発音をサンプリング周波数 44.1kHz、量子化精度 16bit でデジタル化し、音声ファイルとして格納した。後に 6 人の中国語母語話者(女性 3 人, 男性 3 人; 22~33 歳, 平均 25 歳)がその音声を声調で対立する 4 種類のピンインのうちのどれに聴こえるか、前項の知覚実験と同じ強制選択課題を用いて評定した。一人の評定者は 4,320 全音声ファイル(144 刺激語×30 人)を話者(実験参加者)ごとのブロックに分けて評定した。

3. 結果

知覚実験について、参加者ごと声調ごとの平均正答率を求めた。声調(一声～四声)を要因とし、平均正答率を従属変数として F 検定を行った結果、声調によって正答率は有意に変動した[F (3,116) = 12.681, $p < 0.0001$](図 2 左)。一声～四声の正答率はそれぞれ 74.6%, 58.4%, 96.8%, 67.9%であった。

生成では、各発話音声について 6 人の評定者のうち刺激語と同じ語を選択した割合を求めた。本研究ではこの割合を発音明瞭度と呼ぶ。声調(一声～四声)を要因とし、発音明瞭度を従属変数として F 検定を行った結果、声調(一声～四声)の効果は有意だった[F (3,116) = 29.270, $p < 0.0001$](図 2 右)。一声～四声の正答率はそれぞれ 96.3%, 87%, 78%, 98.8%であった。

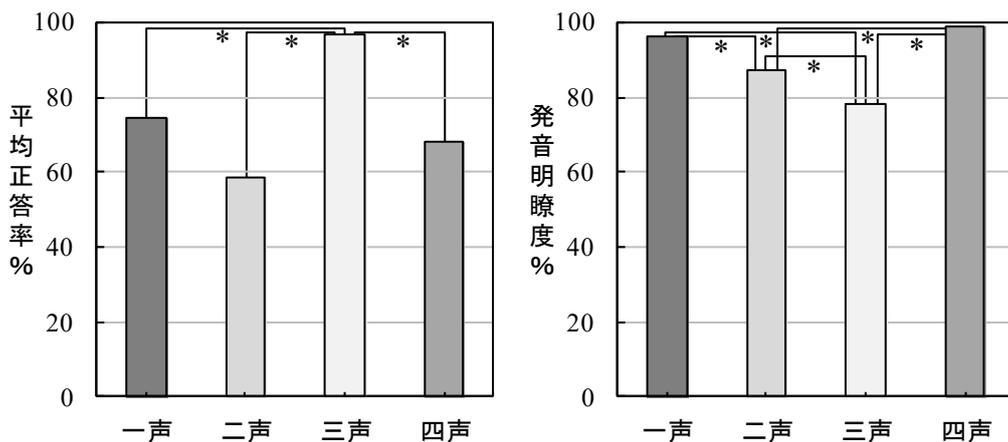


図 2: 知覚実験(左)と生成(右)の結果 * $p < 0.001$

4. 考察

知覚実験の結果では、平均正答率は三声が一声、二声、四声より有意に高かった(三声 > 一声, 二声, 四声). 生成の明瞭度は三声が一声、二声、四声より有意に低く、四声、一声 > 二声 > 三声の順に低下した. 知覚と生成の結果を比較すると、三声は知覚の正答率は最も高いが発音明瞭度は最も低く、一声と四声は知覚の正答率は低いが発音の明瞭度は 100% に近かった. 知覚は容易だが発音が困難、またはその逆の音が存在することが示唆されたといえる.

また、知覚の結果は Wang *et al.* (1999) の知覚実験の結果とは、全体的な傾向は類似していたが、本研究では三声の正答率が顕著に高かった (図 3 左). 生成の結果は Wang *et al.* (2003) のアメリカ英語母語話者の生成の難易度と全体的な傾向は類似していたが、本研究では声調間の差が小さかった.

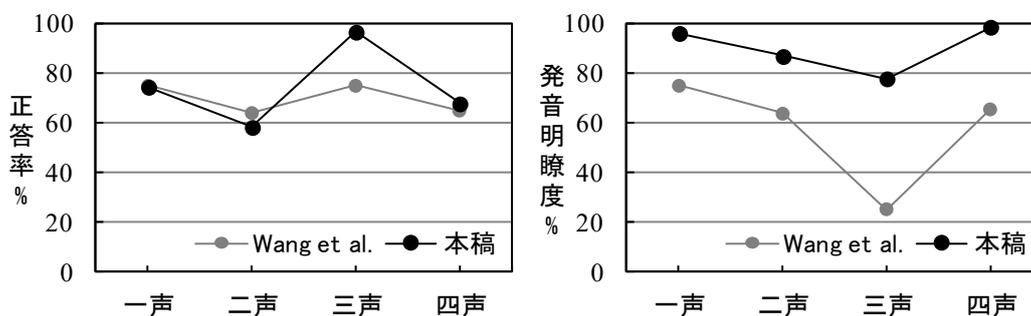


図 3: Wang 等によるアメリカ英語母語話者の結果と日本語母語話者を対象とした本研究結果の比較. 左が知覚の比較 (Wang *et al.*, 1999), 右が生成の比較 (Wang *et al.*, 2003).

Wang 等と本研究と比較すると、課題が多少異なるので直接の比較はできないものの、アメリカ母語話者と日本語母語話者で声調の効果について生成、知覚とも類似した傾向を示したことは、声調の難易度は両母語話者に共通した傾向である可能性を示した。また、本研究では中国語を全く知らない日本語母語話者を対象としたが、特に生成について高スコアを示したことは興味深い。声調のみならず中国語音声に関する知識を持たないは、ピッチの変動を真似することで一時的に明瞭度が上昇している可能性がある。母語である日本語がピッチアクセントであることも関係しているかもしれない。今後、中国語学習過程との関連を検討することにより、効果的な発音習得方法について検討したい。

参考文献

- Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press.
- Chuang, C. K., & Hiki, S. (1972) "Acoustical features and perceptual cues of the four tones of standard colloquial Chinese", *The Journal of the Acoustical Society of America* 52(1A), 146-146.
- Kiriloff, C. (1969) "On the auditory perception of tones in Mandarin", *Phonetica* 20(2-4), 63-67.
- Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999) "Training American listeners to perceive Mandarin tones", *The Journal of the Acoustical Society of America* 106(6), 3649-3658.
- Wang, Y., Jongman, A., & Sereno, J. A. (2003) "Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training", *The Journal of the Acoustical Society of America* 113(2), 1033-1043.
- Yang, S., & Akahane-Yamada, R. (2016) "Effects of perception and vocabulary training of Mandarin tones for native speakers of Japanese: Pre-, post-, and retention test comparison", *The Journal of the Acoustical Society of America* 140(4), 3392-3392.