

残響下における日本語長母音のカテゴリー知覚 —母語話者と非母語話者の比較—*

☆大澤恵里（上智大院・理工研），荒井隆行（上智大・理工），
北原真冬（上智大・外国語）

1 はじめに

残響は大きなホールのような広い空間において音楽の聞こえを豊かにする一方，言語音の明瞭度は低下させる．その要因として self-masking ならびに overlap-masking の影響が指摘されている[1]．後者は先行する音声に伴う残響が後続する音声をマスキングすることによって音声を聞き取りにくくする作用があるが，それと同時に前者によって時間構造が崩されることも音声の明瞭度を下げる要因となっている．

日本語には，例えば「おばさん」と「おばあさん」や「着て」と「切手」からもわかるように，母音や子音の長短によって語の意味が変化する長母音や促音がある．日本語母語話者がこれらの長短を知覚する際，母音や子音の時間長が重要な役割を果たしていることが報告されており，音声の時間構造が重要とされる[2, 3]．また日本語母語話者の長短知覚はカテゴリー知覚と呼ばれ，長短のカテゴリーに基づいて行っていると考えられている．短母音と長母音を時間長の連続体として母語話者に提示し，長短を判別してもらう実験では，「長」「短」の回答が連続体のある箇所で急激に変化し，カテゴリー境界が存在していることを示していた[4]．

残響下での日本語の長短対立の知覚について調べた研究では，残響が母音や子音の時間長の知覚に影響を与えることで，日本語を母語とするリスナーでも長短判別が静かな（残響のない）環境に比べ正確にできなくなることが報告されている[5]．また日本語を母語としないリスナーの知覚に関する研究では，日本語の長短対立の知覚に残響が大きく影響し音声の時間長の変化を残響下ではほとんど聞き取れていないことが報告されている[6]．

残響下での日本語音声の知覚に関する先行研究[5,6]では，母音の長短判別について限られたパターンでしか調査していない．長短判別には母音の物理的な長さだけでなく，話速等のその他の要素も関わっているため，さまざまな条件の母音の長短判別について調査する必要がある．

そこで，本研究では残響下での長短母音判別に焦点を当て，長母音位置と話速を変化させた母音の時間長知覚を残響の有無の2条件で聴取実験を行い調査した．

2 聴取実験

2.1 音声資料

本研究では，母音長で対立するミニマルペアを用いて実験を行った．母音環境を統一するため第一モーラと第二モーラの子音を統一し，母音も/a/に限定した．長母音位置は語頭と語末で，具体的には以下のとおりである．

語頭位置：/baba/ - /baaba/

語末位置：/baba/ - /babaa/

話速は fast, normal, slow の3種類で，fast は8モーラ/秒，normal は6モーラ/秒，slow は4モーラ/秒であった．なお，アクセント型は頭高型に統一した．

音声資料として，20代の東京方言話者（女性）の発話を使用した．発話はキャリア文の「これから聞こえてくるのは___です」に入れたものと単語単体で発話したものを録音した．

2.2 刺激音声

カテゴリー知覚に対する残響の影響を調査するため長短母音の連続体を作成した．長母音語として単語単体で発話された音声を使用し，長母音の周期を決められた周期ずつ切り取ることにより等間隔の10段階で短母音へと変化する連続体を作成した．

なお，本研究ではキャリア文からの

* Reverberation degrades performance of perception of Japanese vowel length contrast by native listeners and non-native listeners, by OSAWA, Eri, ARAI, Takayuki, KITAHARA Mafuyu (Sophia Univ.).

overlap-masking の影響を最小限にするため、刺激音声は単語単体で参加者に提示した。

2.3 参加者

聴取実験には日本語を母語とするリスナー (JL) と日本語を母語としないリスナーが参加した。JL は 20 代の健聴者 6 名であった。日本語を母語としないリスナーとしては英語を母語とするリスナー (EL) と中国語を母語とするリスナー (CL) が参加した。それぞれ 20 代の健聴者が 6 名ずつ参加した。健聴か否かに関しては参加者の自己申告であった。本研究では参加者の母語である英語と中国語の方言は限定しなかった。

2.4 実験方法

実験は上智大学荒井研究室の防音室で行った。防音室に設置された PC 上で音声研究用ソフトウェア Praat[7]を用いて行った。刺激音声を再生した後に PC 画面上に母音長で対立する無意味語を対で表示し、聞こえた音声に対して二者択一で答えてもらった。また、長短の回答と合わせてその回答に対する自信度も 5 段階スケール (5: 自信あり, 1: 自信なし) で評価してもらった。なお、本実験に入る前に話速に慣れるための練習を行ってもらった。練習では本実験で使用する刺激音声をキャリア文に入れて提示した。それぞれのセッションの前に練習を行い、直後に単語単体での本実験を行った。残響条件はすべてのリスナーが残響あり条件のあとに残響なし条件の順番で実験を行った。話速は fast - normal - slow と slow - normal - fast の 2 パターンの順番で行い、この 2 パターンは参加者間でカウンターバランスを取った。

刺激音声は実験に用いた PC からオーディオインターフェース (RME Fireface 800) を介してデジタルミキシングエンジン (Yamaha DME24N) に送られた。同機にはスピーカー (Genelec 8020A) が 4 台接続されており、参加者は 4 台のスピーカーの中央に着席した。各スピーカーから受聴位置までの距離は約 1.3m であった。デジタルミキシングエンジンにはデジタルリバーブ (Roland RSS-303) が接続されており、このリバーブによって刺激音声に擬似的に残響を付加した。付加した残響について、swept sine 信号を用いて音場における受聴位置でのインパルス応答を計測した結果、残響時間は 1.2 秒であった (500-

2kHz における 1/1 オクターブ毎の分析結果の平均値)。なお、残響なし条件はデジタルミキシングエンジンの設定によりデジタルリバーブをバイパスすることにより実現した。この場合の受聴位置でのインパルス応答を計測した結果、残響時間は 0.1 秒であった。

3 分析方法

回答からそれぞれの参加者のカテゴリ知識をモデル化するため、本研究では二項ロジスティック回帰分析を用いた。従属変数は参加者の長短の回答で、独立変数は 10 段階で変化する刺激のステップ番号である。近似曲線を得るため次の数式を用いた。

$$y = \frac{1}{1 + \exp(-a - bx)}$$

x は刺激のステップ番号を表し、 y はその刺激に対する「長」回答の割合を表している。 a と b は二項ロジスティック回帰分析から得られた係数である。

4 結果および考察

4.1 判断の安定性

Table 1 に連続体の両端にあたる刺激 (母音長が最短の刺激と最長の刺激) に対する「長」回答の割合を示す。

Table 1: 母音長が最短と最長の刺激に対する「長」回答の割合 (%)。

			残響なし	
			最短	最長
JL	fast	語頭	0.5%	100.0%
		語末	0.2%	99.8%
	normal	語頭	0.0%	100.0%
		語末	0.1%	100.0%
	slow	語頭	0.1%	100.0%
		語末	0.5%	100.0%
EL	fast	語頭	0.0%	85.2%
		語末	5.4%	96.2%
	normal	語頭	10.0%	99.7%
		語末	1.7%	99.9%
	slow	語頭	1.8%	99.9%
		語末	5.5%	97.1%
CL	fast	語頭	2.0%	99.8%
		語末	28.3%	98.7%
	normal	語頭	1.8%	100.0%
		語末	31.1%	98.5%
	slow	語頭	3.6%	98.3%
		語末	15.5%	98.8%

			残響あり	
			最短	最長
JL	fast	語頭	1.7%	99.4%
		語末	8.4%	93.2%
	normal	語頭	0.1%	100.0%
		語末	0.0%	96.8%
	slow	語頭	0.2%	100.0%
		語末	0.0%	99.8%
EL	fast	語頭	0.5%	97.8%
		語末	12.8%	84.0%
	normal	語頭	16.2%	99.4%
		語末	0.4%	96.6%
	slow	語頭	3.0%	100.0%
		語末	0.9%	99.3%
CL	fast	語頭	3.1%	98.4%
		語末	32.3%	89.1%
	normal	語頭	0.0%	100.0%
		語末	17.1%	99.5%
	slow	語頭	3.3%	100.0%
		語末	23.6%	98.8%

残響なし条件とあり条件の結果を比べてみると、まず JL は残響なし条件では最短の刺激に対する「長」回答の割合がどの条件でもほとんど 0%に等しいが、残響あり条件では母音長が最短の刺激でも話速が fast で語末にある場合は「長」と回答する割合が増加した。また同条件では最長の刺激に対する回答も「長」回答の割合が 93.2%に低下した。話速が normal でも同様の低下が見られ、残響下では語末の長母音の判別が難しくなることがわかる。

EL も JL と同様に、話速が fast の語末母音の知覚では、最短の刺激に対する「長」回答の割合が残響の有無により 5.4%から約 13%に増加し、最長の刺激に対する「長」回答は 96.2%から 84%に低下した。

CL は残響下において、特に語末位置の長母音知覚の安定性が他のリスナーに比べ低かった。話速が fast の場合には最短の刺激に対する「長」回答の割合が 32.3%であり、slow の場合でも 20%以上の割合で「長」と回答していた。

長母音位置別ではどのリスナーも残響下での語末の長短判別が安定しておらず、JL であっても残響下では「長」回答の割合に変化が見られた。本研究で行った聴取実験では刺激音をキャリア文に入れずに提示したため、明確に単語の終わりが定まらず回答が不安定になったと考えられる。

話速別ではどのリスナーも fast の話速で提示した場合で最も回答が不安定になった。ど

のリスナーも最長の刺激に対する「長」回答が低下しており、速度の影響で物理的な母音長が短くなった影響が考えられる。

4.2 長短判別の境界位置

長短カテゴリーの境界位置を「長」回答の割合が 50% になる位置と定めたとき、残響がその境界位置に及ぼす影響を調べた。Fig. 1～Fig. 3 は長母音が語末にある刺激を normal の話速で提示したときの各リスナーの結果である。normal の話速での語末の母音判別がすべてのリスナーにおいて残響の有無によるカテゴリー境界位置の変化が最も大きい条件だった。Fig. 1～Fig. 3 の縦軸は「長」回答の割合を示し、横軸は母音長が変化する刺激のステップを示している。

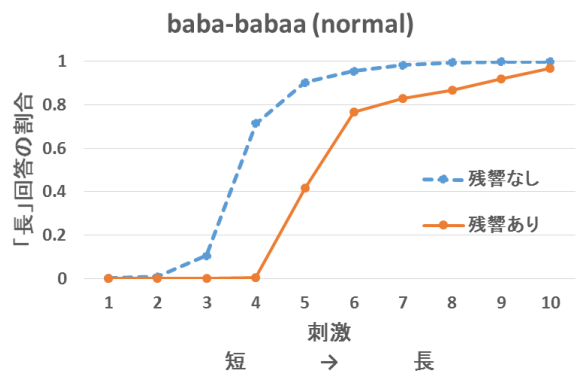


Fig. 1: JL の normal の話速時の /baba/ - /babaa/ 連続体に対する「長」回答の割合。

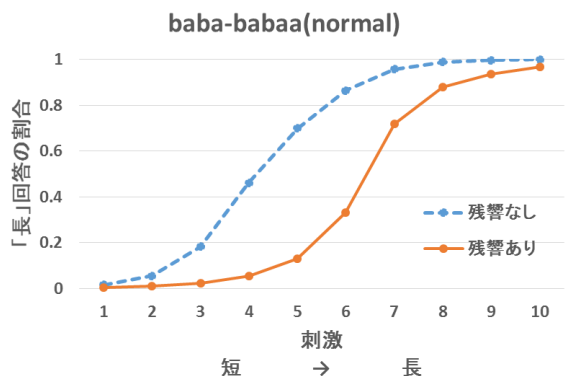


Fig. 2: EL の normal の話速時の /baba/ - /babaa/ 連続体に対する「長」回答の割合。

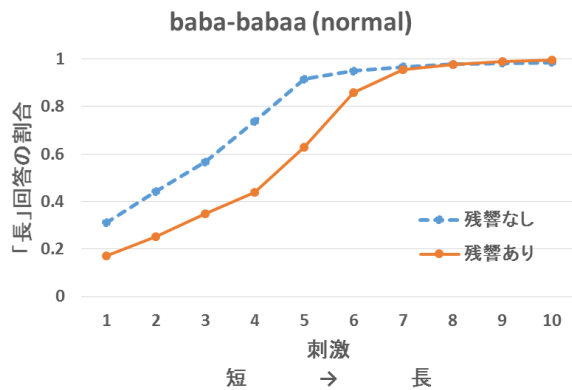


Fig. 3: CL の normal の話速時の /baba/ - /babaa/ 連続体に対する「長」回答の割合.

全てのリスナーに共通して残響あり条件では長短のカテゴリー境界位置が母音長が長い刺激の方向へシフトしている。JLは残響なしの場合カテゴリー境界が刺激のステップの3番目と4番目の間にあったが、残響あり条件では5番目まで移動している。ELの境界位置も刺激のステップを2つ分JLの結果と同じの方向へシフトしている。CLも同様に2ステップ分同方向へ移動しているが、他のリスナーに比べ境界位置が母音長の短い刺激にある。

本稿の聴取実験は、fast – normal – slow と slow – normal – fast の順番を用い、参加者間でカウンターバランスを取った。どちらの順番で行っても normal が2番目に行われていることから、その前に行ったセッションでの残響を normal の実験の際に考慮している可能性がある。感覚として「長」に聞こえたとしても、残響の影響を考慮し「短」と回答する過剰適応が起こっていると考えられることができる。

4.3 参加者の母語別比較

4.2のFig.1～Fig.3を見ると、JLとELの結果では刺激のステップに沿って「長」回答の割合が増加していることがわかり、その割合が変化するステップの位置もほぼ一致していることがわかる。アメリカ英語には母音の長短対立はないが、緊張母音と弛緩母音の対立がある。母音判別の際にはフォルマント等の情報が第一義的な手がかりとされるが、母音長も二義的な手がかりとして用いられている[8]。今回の結果からJLと同様の方法

で日本語の長短判別をしていることが示唆され、母語での母音判別方法が影響している可能性がある。

一方CLの結果では刺激ステップの1番目に対する「長」回答の割合がJLとELに比べ高く、短母音を正確に判別できていない。この傾向は先行研究でも指摘されており、中国語を母語とする日本語学習者を対象にした知覚研究で、短い母音長でも長母音と知覚する傾向が報告されている[9]。また標準中国語には長短の母音対立はない。緊張母音と弛緩母音はあるが、英語とは異なり意味の違いを生じさせない[10]。ELとの結果の差は母語の影響が考えられる。

5 おわりに

本稿では、日本語の長短母音知覚に対し残響の有無が与える影響について調査した。調査対象である長母音は単語内の位置（語頭と語末）と話速を変化させた。残響下においてCLはJLやELに比べ長短判別が不安定になり、母音長が最短と最長の刺激でも判別が安定しなかった。日本語の非母語話者間の聞き取りの差はそれぞれの母語が影響している可能性があった。

参考文献

- [1] A. K. Nábělek, T. R. Letowski, and F. M. Tucker, *J. Acoust. Soc. Am.*, 86(4), 1259-1265, 1989.
- [2] H. Fujisaki, K. Nakamura, and T. Imoto. *Auditory Analysis and Perception of Speech.*, 197-219, 1975.
- [3] 藤崎博也, 杉藤美代子, 岩波講座日本語5, 音韻, 63-98, 1975.
- [4] B. Dawn, T. Arai, P. Czigler, and K. Sullivan, *Proceedings of ICPHS*, 857-860, 1999.
- [5] 荒井隆行, 大澤恵里, 程島奈緒, 井下田貴子, 日本音響学会 2016 年秋季研究発表会予稿集, 293-296. 2016.
- [6] E. Osawa, T. Arai, N. Hodoshima, and T. Igeta, 5th Joint Meeting of ASA and ASJ, poster session, 2016.
- [7] P. Boersma, *Glott International*, 5(9/10), 341-345, 2001.
- [8] R. McAllister, J. E. Flege, and T. Piske, *Journal of Phonetics*, 30(2), 229-258, 2002.
- [9] 栗原通世, 言語科学論集(8), 1-12, 2004.
- [10] Y-H. Lin, *The sounds of Chinese*, 2007.