

入力モダリティによる音韻表象のちがい： 促音・非促音の音韻カテゴリ境界を手掛かりとした検討

Difference of phonological representation between auditory and visual language processing using geminate/non-geminate categorical boundary

中村 暢佑[†], 麦谷 綾子[‡], 渡邊 淳司^{†‡}
Yosuke Nakamura, Ryoko Mugitani, Junji Watanabe

[†] 東京工業大学大学院 総合理工学研究科,

[‡] 日本電信電話株式会社 コミュニケーション科学基礎研究所

[†] Inter disciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology

[‡] NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation

nakamura.y.bj@m.titech.ac.jp

mugitani.ryoko@lab.ntt.co.jp

watanabe.junji@lab.ntt.co.jp

Abstract

Humans have the language processing mechanisms involving auditory perception of speech and visual perception of letters in the brain. The auditory processing requires phonological analysis of speech whereas the visual processing primarily requires spatial morphological analysis of letters. However, a previous study indicated the processing of analyzing visual stimuli (letters) also activates brain areas that are responsible for phonological processing. Here, we studied whether the phonological representations elicited by auditory and visual language processing are the same or not. Using geminate/non-geminate category classification task, boundary shift of phonemic category was compared among different modality conditions. The result revealed phonemic boundary shift corresponding to the input modality of the standard stimulus, and hence supported the idea that phonological representation is specific to the modality.

Keywords — Language, Word comprehension, Geminate word, Phonological representation

1. はじめに

私たちの脳には、大きく分けて、音声に関する聴覚由来の言語処理機構と、文字等に関する視覚由来の言語処理機構の2つが存在する[1]。聴覚に関する言語処理機構は側頭葉がつかさどり、時系列的に提示される音韻から単語や文章の理解を行うものである。一方、視覚に関する言語処理は、頭頂葉がつかさどり、主に、文字等の空間パターンを分析し、それらが指し示す意味を理解するも

のである。ただし、これまでの研究で、文字を見た時にも、その文字の音韻を心の中で読み上げるような処理が半自動的に行われることが示されている[2][3]。つまり、視覚的な言語処理においても、常に音韻処理が行われるということである。本研究では、聴覚的な言語処理における音韻表象と、視覚的な言語処理において生起される音韻表象が、同じ性質を示すかどうかについて検討した。具体的には、音声で提示される言語情報に対する音韻処理と、文字で提示される言語情報に対する音韻処理と、さらには、絵の図像情報の意味を介した音韻処理について比較検討した。文字と絵は、両者とも初期視覚野からの情報をもとにするが、文字、特に表音文字は初期視覚野のある後頭葉から直接頭頂葉へ情報処理が進む一方で、絵は後頭葉から側頭葉でパターン認知が行われる。そのため、同じ視覚情報の文字と絵でも、必ずしも同じ音韻処理が行われるとは限らない。

実験では、「バター」もしくは「バッター」という意味を持つ刺激（基準刺激）を、音声（/bata:/と/batta:/）、文字（「バター」と「バッター」という文字）、または絵（バットを持つバッターの絵と皿の上のバターの絵）のいずれかで提示した。その後、提示する音声（ターゲット刺激）が基準刺激と同じ語のカテゴリ（「バター」もしくは「バッ

ター)」であるかを回答してもらった。「バター」と「バター」の音声は促音（「っ」）の有無のみが異なり、非促音（バター）・促音（バター）は、日本語母語成人にはカテゴリーカルに知覚される。そのため、ある一定の無音区間長を境に、それよりも短い無音区間を持つものは非促音、長い無音区間を持つものは促音として判定される[4]。無音区間の長さを操作し、「バター」から「バター」へと段階的に変化する刺激群をターゲット刺激として用い、「バター」と「バター」の知覚頻度が50%となる時間間隔をカテゴリー境界とした。基準刺激の入力モダリティ（聴覚 v.s. 視覚）や音韻互換性（文字 v.s. 絵）の違いによって、ターゲット刺激のカテゴリー境界がどの程度変化するのか検討した。このタスクにおいて、被験者は、基準刺激に対する音韻表象をもとにターゲット刺激のカテゴリーを判断することになる。どの基準刺激に対してもターゲット刺激は同じ音声刺激を用いているため、もし、ターゲット刺激のカテゴリー判断の境界が異なれば、基準刺激の音韻処理が異なるということになる。

2. 実験条件

2.1. 被験者

実験に参加した被験者は満 22~38 歳までの男女 10 名（男性 4 名、女性 6 名）である。本実験は NTT コミュニケーション科学基礎研究所の倫理委員会によって承認され、全ての被験者から実験の参加について同意を得た。

2.2. 実験刺激

実験に使用した音声は、30 代女性のプロナレーターが発声した「バター」の音声を、防音室内においてサンプリングレート 44.1kHz で収録した。その後、/ba/と/ta/の音節間の無音区間に人為的に 10ms ずつの無音を挿入することで、「バター」から「バター」まで段階的に分類される 9 種類の音声（無音区間長：130ms-210ms）を作成し、ターゲット刺激として使用した。基準刺激として、「バター」と「バター」に対応する音声、文字、

絵をそれぞれ用意した。音声刺激については、無音区間長が 130ms の音声を「バター」、210ms の音声を「バター」として使用した。視覚刺激については、「バター」と「バター」の文字が描かれた画像ファイルを文字刺激として使用し、皿の上のバターとバットを持つバターの絵が描かれた画像ファイルを絵の刺激として、それぞれ使用した。実験は暗室で行った。すべての聴覚刺激は、ノート PC の組み込みスピーカーを使用して提示した。また、すべての視覚刺激は、ノート PC のモニタ上に提示した。被験者は、ヘッドホンなど、特別な装置を装着することはなかった。回答方法については、ノート PC の組み込みキーボードの矢印キーを、「←」を「NO」、「→」を「YES」に割り当て、聴取した音声の回答に使用した。

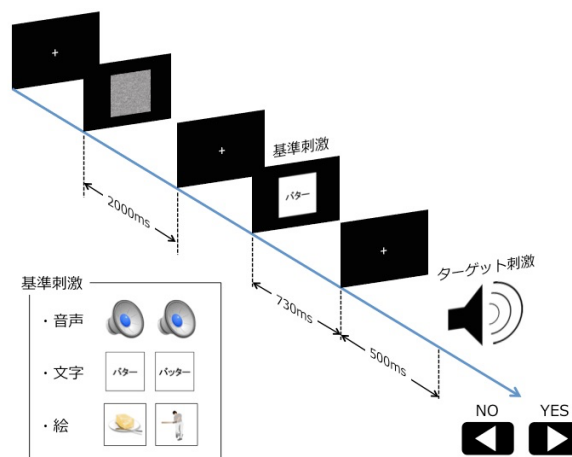


図 1 実験手順（基準刺激が文字の場合）

2.3. 実験手順

実験の各ブロックで提示される基準刺激は、音声、文字、絵のいずれか一種類であった。実験の各試行においては、「バター」もしくは「バター」のいずれかがランダムに提示され、その 500ms 後に提示された音声（ターゲット刺激 9 種類のうちひとつ）のカテゴリが前に提示された基準刺激のカテゴリと同じかどうかを回答してもらった（図 1）。各試行の間には、前の試行の影響を除去するために、基準刺激が音声の場合にはビーブを、文字や絵の場合にはスノーノイズを提示した。音声、文字、絵の各ブロックは基準刺激 2 種類に対してターゲット刺激 9 種類それぞれを 5 回実施し

(90 試行), それぞれ 5 ブロック実験行っった. 各ブロックの実施順序は被験者間でカウンターバランスをとった.

3. 実験結果

各刺激条件において, 「バター」と「バッター」それぞれが基準刺激であった場合のカテゴリ境界となる無音区間長を, 被験者ごとに求めた. 具体的には, 各実験の結果から, 2 種類の基準刺激(「バター」もしくは「バッター」と 9 種類のターゲット刺激の組み合わせのそれぞれに対して, ターゲット刺激が基準刺激と同じであると回答した確率を算出し, 最尤推定法による心理測定関数の最適当てはめを行った(図 2). このとき, 「基準刺激と同じである」と回答した確率が 50%となる無音区間長をカテゴリ境界とし, それを基準刺激ごとに求めた(図 3). そのカテゴリ境界の値について, 被験者内 2 要因(要因 A: 基準刺激のカテゴリ(2 水準: バター, バッター) × 要因 B: 基準刺激(3 水準: 音声, 文字, 絵))の分散分析を行った. その結果, 要因 A に関して, 基準刺激のカテゴリが「バッター」の場合に比べて, 「バター」の場合のカテゴリ境界の値が有意に大きかった($F(1, 9)=5.290, p<.05$). すなわち, 基準刺激が「バター」の場合には知覚的に曖昧なターゲット刺激を「バター」と判定しやすく, 一方, 基準刺激が「バッター」の場合には, 同じターゲット刺激を「バッター」と判定しやすい. このことは, 基準刺激によって, ターゲット刺激のカテゴリ境界値がシフトすることを示している. さらに, 要因 B に関して, 基準刺激によってカテゴリ境界の値が異なることが示された($F(1, 9)=9.885, p<.005$). 下位検定から, 基準刺激のモダリティが音声であると, 視覚情報である文字や絵に比べてカテゴリ境界の値が有意に大きくなった(音声-文字: $t(18)=3.951, p<.001$, 音声-絵: $t(18)=3.742, p<.005$). つまり, 同じ無音区間長ならば, 音声のほうが「バター」と判定しやすいということである. 一方, 要因 A と B に関して, 有意な交互作用はみられなかった.

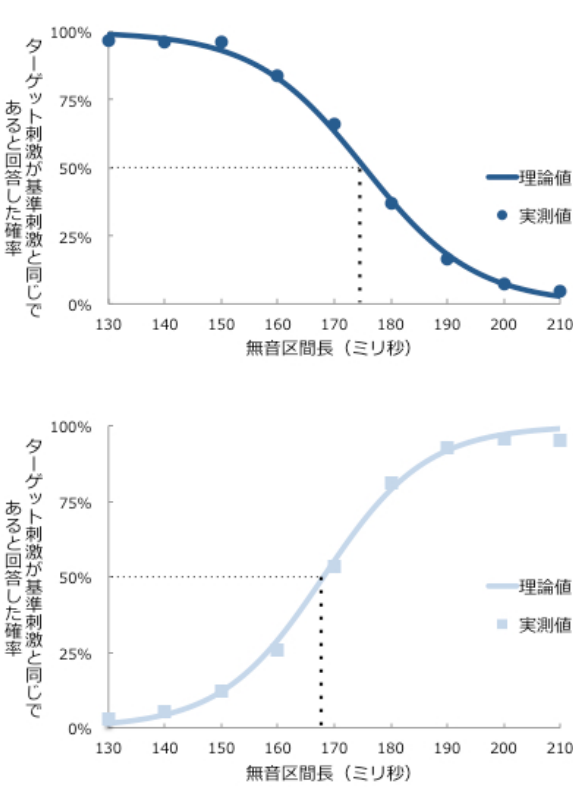


図 2 ターゲット刺激が基準刺激と同じであると回答した確率(点線はカテゴリ境界を示す)
基準刺激: /bata:/ (上), /batta:/ (下)

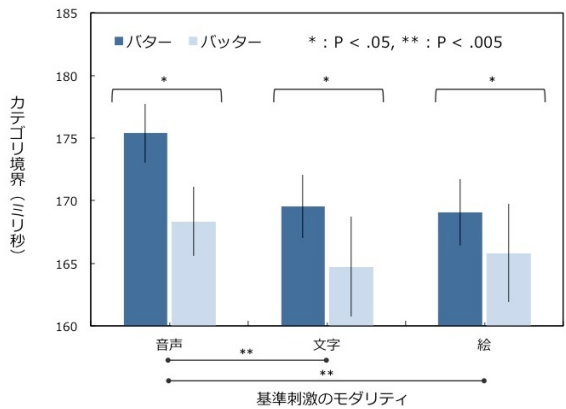


図 3 基準刺激の各モダリティにおけるカテゴリ境界値と標準誤差

4. 考察とむすび

本研究では, 入力モダリティが異なるときに, そこで生起する音韻表象が異なるかどうか検証する実験を行った. その結果, 直前に提示された基準刺激のカテゴリによって, 無音区間をもつ音声のカテゴリ判断は影響を受けることがわかった.

具体的には、基準刺激が「バター」の場合には、「バター」と判断される無音区間長の領域が増加し、「バッター」の場合には、「バッター」と判断される無音区間長の領域が増加した。つまり、聴取する基準刺激の方向にカテゴリ境界値がシフトした。また、文字や絵の基準刺激を視覚提示した場合に比べて、音声を聴覚提示した場合にカテゴリ境界の値が有意に大きくなったことから、刺激の知覚モダリティによって異なる音韻表象が生成される可能性がある。一方、文字と絵は音韻への互換性に違いがあるにもかかわらず、境界が変化しなかったことから、同じ視覚への入力であれば、音韻の互換性によらず類似した表象が生成される可能性が示唆される。一方で、本実験ではターゲット刺激として音声を用いたため、基準刺激が音声の場合は基準刺激とターゲット刺激は同じモダリティに提示され、文字や絵の場合は違うモダリティに提示された。したがって、基準刺激が聴覚提示された場合にカテゴリ境界値が大きくなる現象は、聴覚・視覚といったモダリティの違いによって生じるのではなく、基準刺激とターゲット刺激のモダリティが同じことによって生じた可能性は否定できない。今後、この点を検証できるような課題を設定し、そのメカニズムを解明していきたい。

参考文献

- [1] Cathy J. Price, (August 1998) “The functional anatomy of word comprehension and production”, Trends in Cognitive Science, Vol. 2, No. 8, pp. 281-288.
- [2] Price, C. J., Wise, R.J.S. and Frackowiak, R.S.J. (1996) “Demonstrating the implicit processing of visually presented words and pseudowords”, Cerebral Cortex, Vol. 6, pp. 62-70.
- [3] Macleod, C.M. (1991) “Half a century of research on the Stroop effect: an integrative review”, Psychological Bulletin, Vol. 109, pp. 163-203.

- [4] Fujisaki, H., Nakamura, K. and Imoto, T., (1973) “Auditory perception of duration of speech and non-speech stimuli”, Annual Bulletin RILP, Vol. 7, pp. 45-64.