

## 日本語における促音の知覚とモダリティ\*

○荒井隆行（上智大・理工）

## 1 はじめに

日本語における促音では、音響的に子音の持続時間が長くなることが知られている[例えば 1, 2]. 特に無声破裂子音の場合、閉鎖区間が伸長する. 一方、先行研究[3]では、/atta/と/a+pause+ta/を比較している. そして、先行母音の末尾以外、音響的にはまったく同じ/atta/を合成し、先行母音の末尾にフォルマント遷移がある刺激とない刺激を使って知覚実験を行った. その結果、/t/へのフォルマント遷移がある場合は/atta/の回答率が上昇する一方、フォルマントが遷移せず一定のまま無音区間（閉鎖区間）に入るような刺激では、/a+pause+ta/の回答率が上昇した. このことから、日本語母語話者にとって促音に聞こえるためには先行母音の末尾にフォルマント遷移があることが知覚上、重要であることがわかった.

ところで、私たちの音声コミュニケーションでは視覚情報も重要であり、様々な研究がそれを示している. 先行研究[4]では、/a+pause+ta/の合成音声に同期するように発話された/atta/の母語話者の映像を組み合わせるなどして、知覚実験を行った. その結果、視覚情報がない場合は/a+pause+ta/の回答が多かった音声に対し、同じ音声に/atta/の映像が組み合わせることで/atta/回答率が上昇した. なお、ここで/atta/と/a+pause+ta/の映像では閉鎖が開始する時点が早いか遅いかだけでそれ以外はほとんど差がなかった. 映像が音声知覚に与える影響では、McGurk 効果[5]が有名であるが、この研究[4]においても視覚による調音ジェスチャーが音声知覚を補う効果を示す例と言える.

このようにマルチモーダルな音声知覚の研究として、我々はその続きとなる実験結果を報告した[6]. そこでは、もう1つのモダリティとして触覚についても議論している. これは、文献[7-9]などにおいて体性感覚や触覚が音声知覚に影響を与えることが

報告されていることによる. そこで本稿では、我々の先行研究[6]をレビューすると同時に、新たな考察を試みる.

## 2 先行研究[6]について

実験では3つの刺激パターンがあり、聴覚のみ、聴覚と視覚、聴覚と触覚であった.

聴覚刺激は先行研究[4]に準じて/atta/の合成音声を基本とし、母音はフォルマント合成[10]、子音は実際の母語話者の音声を切り出して使用した. 先行母音の末尾のフォルマント遷移は3段階のものを用意し、S0は遷移なし、S1は遷移あり（遷移部の持続時間を50ms）、S2は遷移あり（遷移部の持続時間を80ms）とした.

視覚刺激はFig. 1のように、文献[11]の声道模型の一部を3Dプリンタ出力しそれをロボットハンド（Vstone, Robovie-nano）に装着した. そして、PCからArduinoを介してサーボモータ（Vstone, VS-S020A）によって動作を制御した.

触覚刺激はFig. 2のように、視覚刺激の際に用いたのと同じ制御系でロボットハン

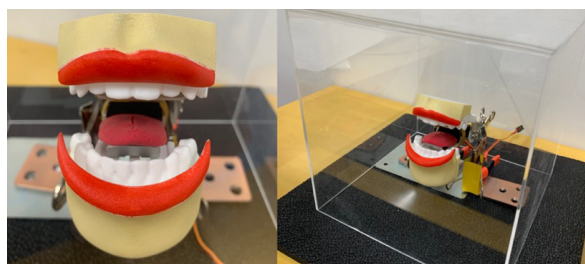


Fig. 1: Physical models of upper/lower jaws and tongue used for visual stimuli [7].

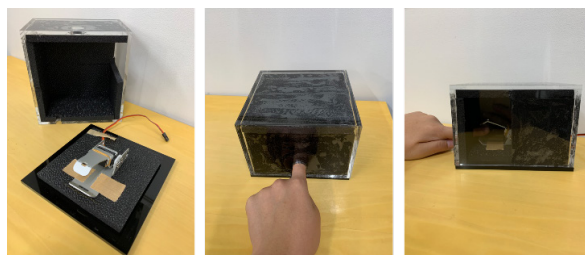


Fig. 2: Same robot-hand unit was used for tactile stimuli [7].

\* Perception of geminate consonants and modality in Japanese, by ARAI, Takayuki, (Sophia University).

ドを動かし、その間に指を置いて挟まれた感覚を与えた。

実験の結果、フォルマント遷移がない場合に聴覚刺激のみでは/a/+pause+/ta/の回答が多かった (/atta/回答率が 10.5%) のに対し、視覚や触覚刺激があることで/atta/回答率が上昇した (聴覚+視覚=40.3%, 聴覚+触覚=39.5%)。このことから、視覚のみならず触覚から得られる動作感覚が聴覚情報を補い、促音の知覚を促したと考えられる。

### 3 考察・まとめ

先行研究[6]に対して、さらなる考察を議論する。まず、[4]と[6]の実験を比較すると、無音区間 (閉鎖区間) の長さが 380 ms から 700 ms と長くなっている点と、視覚刺激については実際の人間の顔映像から声道模型に変わっている点が大きく異なる。それにも関わらず、[4]では/atta/回答率が聴覚刺激のみの 9.3%から聴覚+視覚の 42.7%に上昇しており、[6]も同程度の結果となっている。さらに、聴覚+触覚でも同程度となっているところがポイントである。促音の知覚には先行母音末尾に、何らかの「促音がここから始まります」というキューが必要であることは[12]でも報告しており、それが仮に声門閉鎖であってもその役割を担っていた。触覚刺激から得られる感覚も閉じるという動作であり、同様のキューとなっていたと考えられる。

また、今後の課題としていくつかの点も考えられる。まず、調音位置による違いまで考慮した場合にどうなるかは興味深いところである。一方、フォルマント遷移がない聴覚刺激と組み合わせたマルチモーダル条件で、/atta/回答率が約 40%となっており、これが限界なのか、それとも何らかの理由 (例えば合成音声の質なども含む) かは検討の余地がある。フォルマント遷移部がノイズでマスクされている場合、/atta/回答率が上がる可能性もある。そして、聴覚+視覚+触覚による回答率上昇も期待される。

### 謝辞

内容の一部は、JSPS 科研費 21K02889 ならびに上智大学重点領域研究の助成を得た。

### 参考文献

- [1] M. S. Han, “The feature of duration in Japanese,” *Onsei no Kenkyu*, 10, 65–80, 1962.
- [2] Y. Hirata, “Perception of geminated stops in Japanese word and sentence levels,” *Onsei-gakkai-kaiho*, 194, 23–28, 1990.
- [3] E. Yanagisawa and T. Arai, “Effects of formant transition and intensity damping of preceding vowel off-glide on perception of Japanese geminate consonant Sokuon,” *J. Acoust. Soc. Jpn.*, 71(10), 505–515, 2015.
- [4] T. Arai, E. Iwagami and E. Yanagisawa, “Seeing closing gesture of articulators affects speech perception of geminate consonants,” *J. Acoust. Soc. Am.*, 141(3), EL319–EL325, 2017.
- [5] H. McGurk and J. MacDonald, “Hearing lips and seeing voices,” *Nature*, 264, 746–748, 1976.
- [6] T. Arai, M. Yamada and M. Okusawa, “Syllable sequence of /a/+ta/ can be heard as /atta/ in Japanese with visual or tactile cues,” *Proc. INTERSPEECH*, 3083–3087, 2022.
- [7] S. Tremblay, D. M. Shiller and D. J. Ostry, “Somatosensory basis of speech production,” *Nature*, 423, 866–869, 2003.
- [8] T. Ito, M. Tiede and D. J. Ostry, “Somatosensory function in speech perception,” *Proc. Natl. Acad. Sci.*, 106, 1245–48, 2009.
- [9] B. Gick and D. Derrick, “Aero-tactile integration in speech perception,” *Nature*, 462, 502–504, 2009.
- [10] D. H. Klatt, “The new MIT speech VAX computer facility,” *Speech Communication Group Working Papers IV*, Research Laboratory of Electronics, MIT, Cambridge, 73–82, 1984.
- [11] T. Arai, “Flexible tongue housed in a static model of the vocal tract with jaws, lips and teeth,” *Proc. INTERSPEECH*, 171–172, 2018.
- [12] 荒井隆行, 岩上恵梨, 柳澤絵美, “日本語促音が開始した可能性を知らせる知覚上の手がかり,” 音講論(秋), 297–300, 2016.