

平成24年度研究助成 【サウンド技術振興部門】より

ヒト聴覚の時間、空間、音声の認識に 共通する無音の役割



九州大学大学院
システム情報科学研究院
情報学部門教授

森 周司



九州大学大学院
システム情報科学府
情報学専攻修士課程

小山 和紀



九州大学工学部
電気情報工学科

菊池 洋祐

1. はじめに

無音とはまさに「音が無い」状態であり、私達には何も聞こえない。「音が聴こえる」状態を主な対象とする聴覚研究では、無音が主役となることはない。しかし、主役が名脇役の存在によって引き立つように、音も無音という脇役がいるからこそ際立つと考えることも出来る。この研究は、時間、空間、音声の認識において無音が主役級ではないにしても重要な役割を果たすことを示すものである。

2. 無音の役割

私達は聞いている音が途切れるとすぐ気付く。レコードやCDの音飛びがその良い例である。音飛びの部分を無音と見なすと、これは連続音中の無音検出事態である。無音検出は、音の時間的変化に対する感度を測る方法として良く用いられる。

無音検出では一般に無音前後の音（先行音と後続音）の周波数が同じであり、無音が1～5 msec程度であれば十分に検出可能である。しかし先行音と後続音の周波数が異なる周波数間無音検出では2音の周波数が2オクターブ離れると無音検出閾値は50msec超に達する¹⁾。

何故、周波数間無音検出でこれほど時間分解能が落ちるのかについては、これまで明確な理由は示されていない。一つの有力な説は、Dennis P. Phillipsが提唱するチャンネル間処理である¹⁾。ヒトの聴覚機構は、内耳蝸牛の基底膜から始まる周波数地図を概念化した、複数の周波数チャンネルの集合で表現される。Phillipsによれば、周波数内無音検出は同一周波数チャンネルに流れる音の不連続検出である。一方、周波数間無音検出では先行音と後続音が異なるチャンネルで処理されるため、二つのチャンネルの処理状況を監視する必要が

ある。これがチャンネル間処理であり、聴覚系中枢で行われると考えられる。周波数間無音検出での高い閾値は、このチャンネル間処理の時間分解能を反映するのであろう。

Phillipsはチャンネル間処理が聴空間と音声の認識にも関わることを示唆している。彼のグループは左右の耳をチャンネルと想定して無音の先行音と後続音を左右別々の耳に呈示する両耳間無音検出実験を行い、検出閾値が50msec超になることを示した¹⁾。これは先行音と後続音を別々のスピーカーから出しても同じであり²⁾、聴空間内の位置が異なるチャンネルとして機能していることを示唆する。音声に関しては、有声開始時間（voice onset time、以下VOT）による音声識別が周波数間無音検出と類似していることを指摘している。図1に音声波形中のVOTを示している。VOTはほぼ無音の区間と言ってよく、その前後では周波数成分が全く異なる。*/ba/*と*/pa/*は主にVOTの長さにより識別され、その境界は英語では40～

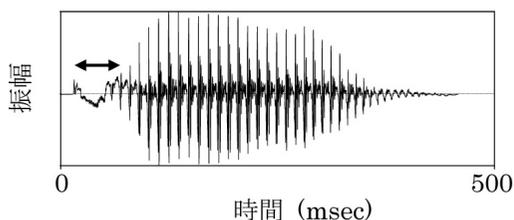


図1 */pa/*の波形中のVOT（矢印部分）

50msecと周波数間無音検出閾値と近い。実際、個人内の識別境界VOT値と周波数間無音検出閾値に有意な正の相関があることが報告されている³⁾。

3. 現在の研究

現在我々は、周波数間及び両耳間無音検出閾値と*/ba/-/pa/*の識別境界VOT値の関係を調べている。これらの値を音響心理学実験で測定するのだが、苦労したのは*/ba/*と*/pa/*の音声刺激の作成である。識別境界VOT値を決定するには、VOT以外の音響特性を出来る限り変えずに、VOTを段階的に変えて*/ba/*から*/pa/*と知覚的に変化する刺激を作成しなければならない。その方法としては、フォルマント合成器を用いて人工音声を生成する、音声を録音して加工する、など色々な方法はあるが、我々自身が聴いて満足するような自然な音声を作成することが出来なかった。そこで、今回はDonders研究所（オランダ）の貞方マキ子研究員の紹介で、同研究所のAlex Brandmeyer研究員が作成した音声刺激⁴⁾を使わせて頂くことにした。図1に示した音声もその一部である。図2にこの音声刺激を用いて我々が行った*/ba/-/pa/*識別実験の結果を示す。VOTが40msec付近で*/ba/*と*/pa/*の判断が入れ替わっており、典型的なカテゴリー知覚を示している。

周波数間無音検出では先行音を中心周波数

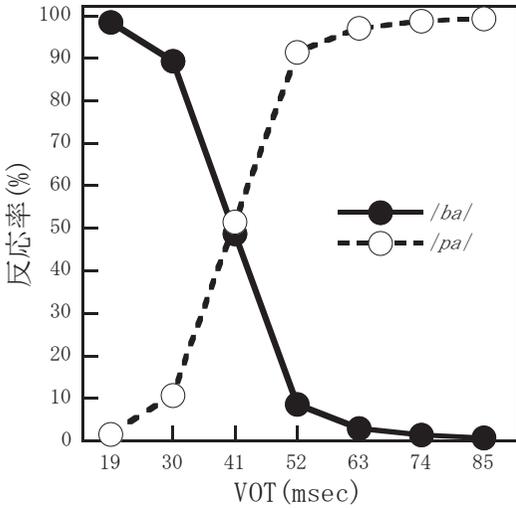


図2 日本語母語者22名の/ba/と/pa/の平均反応率

3505Hzの広帯域雑音、後続音を中心周波数が700、1400、2800Hzの1/2オクターブ帯域雑音とした。両耳間無音検出では無音の先行音と後続音を中心周波数700Hzの1/2オクターブ帯域雑音とし、先行音を左耳、後続音を右耳に呈示した。図3に無音検出閾値と識別境界VOT値の平均を示す。両耳間無音検出閾値と識別境界VOT値は先行研究^{1),3)}の値に近いが、周波数間無音検出閾値は先行研究のそれよりも小さい。各値間の相関を取ったところ、いずれも有意ではなかった。刺激や実験手続きを変えて次の実験を準備中である。

また並行して、/ba/-/pa/音声刺激を子音部と母音部に分け、それぞれを左右別の耳に呈示する音声識別実験を行っている。両耳間無音検出では無音閾値が長くなることを利用し、左右別々の耳に呈示することにより知覚的なVOTが短くなり識別境界もそれに伴って移動すると考えたのである。その途中結果を図4に示す。我々の予想通り、左右別々の耳に呈示した場合に識別境界 (50%反応率点) が長くなっている。

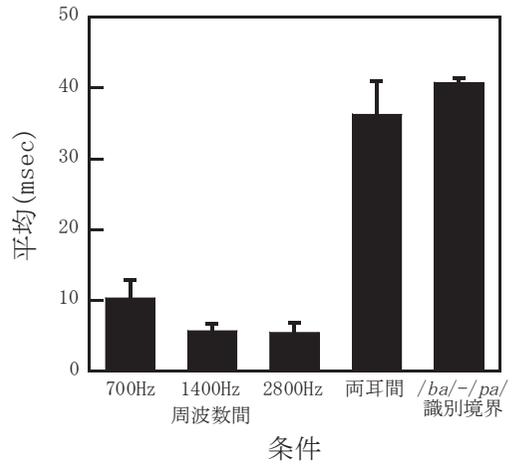


図3 日本語母語者22名の周波数間（後続音中心周波数700、1400、2800Hz）及び両耳間無音検出閾値、識別境界VOT値の平均（エラーバーは標準誤差）

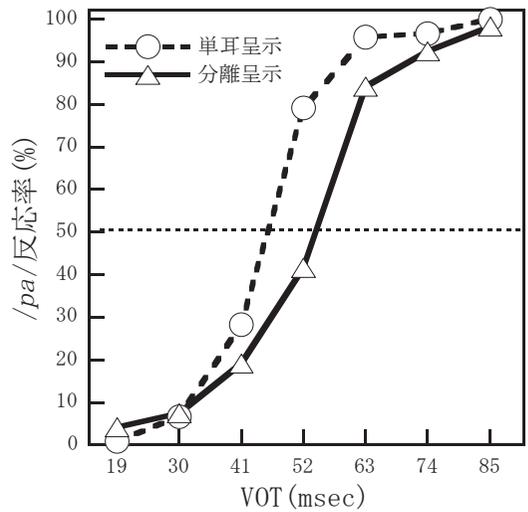


図4 日本語母語者6名の音声単耳呈示と左右耳分離呈示での/pa/平均反応率

但し、この実験では分離呈示した子音部あるいは母音部だけでも/ba/もしくは/pa/と聞こえていた可能性があり、現在はそれを修正した実

験を準備中である。

まだまだ研究途上ではあるが、これまでの研究では先行研究にはない新たな知見も得られており、我々は無音の役割の解明という最終目標に向けて実験と考察を繰り返している。

謝 辞

本研究に対して研究助成を賜りました、一般財団法人カワイサウンド技術・音楽振興財団に心より感謝申し上げます。また音声刺激を提供して頂いたAlex Brandmeyer氏、そして貞方マキ子氏にこの場を借りてお礼申し上げます。

参考文献

- 1) Phillips, D. P. (1999). Auditory gap detection, perceptual channels, and temporal resolution in speech perception. *Journal of the American Academy of Audiology*, **10**, 343 – 354.
- 2) Phillips, D. P., Hall, S. E., Harrington, I. A., & Taylor, T. L. (1998). “Central” auditory gap detection : A spatial case. *Journal of Acoustical Society of America*, **103**, 2064 – 2068.
- 3) Elangovan, S., & Stuart, A. (2008). Natural boundaries in gap detection are related to categorical perception of stop consonants. *Ear & Hearing*, **29**, 761 – 774.
- 4) Brandmeyer, A., Desain, P. W. M., & McQueen, J. M. (2012). Effects of native language on perceptual sensitivity to phonetic cues. *Neuroreport*, **23**, 653 – 657.