

日本語の語頭閉鎖音における有声・無声対立の実現に 発話環境が与える影響 —残響環境と雑音環境の比較—*

☆川井 友子 (上智大院・理工), 李 瑋昱, 辻 慎也, 荒井 隆行 (上智大・理工)

1 はじめに

本研究では雑音下での発話 “Lombard speech,” そして残響下での発話を観察し、発話環境によって発話が変化するか、そして日本語の語頭閉鎖音の有声無声対立においても実現に変化が生じるのかについて調査した。

Clear speech [1]と Lombard speech [2]では、発話音声のインテンシティの増加、話速の低下、平均基本周波数 (f_0) の増加、母音の持続時間の増加、長時間平均スペクトルにおける 1000–3000 Hz の高周波帯域でのエネルギーの増加、変調スペクトルにおける低次の変調周波数帯域のエネルギー増加など共通した音響的变化が報告されている。Clear speech ではそれに加えて母音空間の拡大、 f_0 のレンジの増加なども報告されている [1]。また残響下での発話でも、残響時間などの音場特性の違いによって変化の程度は異なるが、類似した音響的变化が報告されている [3–5]。これらの類似した音響的变化は、発話環境によって生じる生理的反射だけではなく、意図的な発話調整も関与している可能性を示す。

Ishikawa *et al.* (2024) は、通常発話と Clear speech で各発話スタイルを統制した結果、話速とインテンシティは発話スタイル間では変化した一方、静音下と残響下の間における発話変化では一貫した変化が生じなかったことを報告した [6]。そこで本研究では先行研究 [6]と同様に発話スタイルを統制した上で、先行研究より多くの音響指標を対象に、発話環境による変化を調査し、雑音や残響下における発話変化に含まれる生理的反射と意図的な発話調整の寄与、あるいはそれらの間で相乗効果が生じるのか検討した。加えて、発話環境の変化が音韻対立に関わる音響的特徴にも影響を及ぼすかを明らかにするために、日本語の語頭閉鎖音の有声無声対立に焦点を当て、若年の首都圏出身の日本語母語話者を対象に

調査した。

川井他 (2025) は、Clear speech における残響下と静音下での発話を比較した結果、日本語の有声閉鎖音 /b/ と接近音 /w/ の対立に関わる第 1 フォルマント (F1) の遷移の持続時間に変化が生じなかったことを報告した [7]。しかし先行研究 [7]では通常発話での残響の影響や、有意味語による検討は行われておらず、また /b/ と /w/ には語彙的制約があった。そこで本研究では、日本語の語頭閉鎖音の有声・無声対立を対象に、有意味語によるミニマルペアとそれに準じた語のペアを用い、通常発話と Clear speech での残響下と雑音下における発話傾向を観察した。その際に、発話全体での音響的特徴と有声無声対立に関わる音響的特徴である VOT (Voice Onset Time) と閉鎖音の後続母音の onset の f_0 [8]に着目した。また、有声閉鎖音における先行声帯振動は有声無声対立に関する主要な音響的特徴だが、近年では語頭の有声閉鎖音の +VOT 化が報告されている [9]。そこで本研究では、/b/ で正の VOT を示した発話と /p/ の発話を分析対象とし、有声と無声間で VOT の重なりが生じるような場合、発話環境の変化によって VOT における対立が強化されるのか検討した。また、後続母音の onset の f_0 についても同様に検討した。

2 手法

2.1 話者

首都圏出身の日本語母語話者 9 名 (男性 4 名、女性 5 名) のデータを分析対象とした。全話者は関東の首都圏で育ち、0–10 歳までの間に他地域に居住した経験はなかった。話者の年齢は 19–22 歳 ($M = 20.6$ 歳, $SD = 0.9$) であった。

2.2 刺激語彙

発話課題には、語頭にターゲット子音である閉鎖音 /b/, /p/ を含む単語を用い、頭高型の

* Patterns of change in cues for the Japanese word-initial voicing contrast under reverberant and noisy speaking condition, by KAWAI, Tomoko, LI, Weiyu, TSUJI, Shinya and ARAI, Takayuki (Sophia Univ.).

ピッチアクセント (以降 HL と呼ぶ) のミニマルペア「バス」と「パス」、第1モーラが低く第2モーラが高くなるピッチアクセント型 (以降 LH と呼ぶ) におけるミニマルペアに準じた単語ペア「バーテンダー」と「パート」を使用した。加えて、後続母音の onset の *fo* において分析時の基準値を設定するため、語頭に /m/ を含む単語で HL の「迷子」、LH の「舞妓」を刺激語に加えた。これらの刺激語は Gao and Arai (2019) [8] の刺激語彙を参照して選定した。

各試行では、GUI 上に刺激語を表すイラストを提示し、その下部に刺激語を文字でも提示した。参加者には、音声で提示した「これは何ですか？」という質問に対し、キャリア文の「これは、__です。」の空欄に刺激語を挿入して回答するように求めた。また、刺激語が明確な語頭位置で発話されるよう、「これは、」と刺激語の間に十分な間隔を空けて発話するように教示した。

2.3 実験手順

実験は通常発話と Clear speech の2種類の発話スタイル、そして静音下、雑音下、残響下の3つの発話環境を組み合わせた合計6ブロックで構成されていた。通常発話条件では発話スタイルに関する教示は行わず、Clear speech 条件では明瞭に発話するように教示した。発話スタイルごとにセッションを分け、2つのセッションで日を分けて実施した。また、全話者にて必ず通常発話スタイルのセッションから実施した。そして、各発話スタイル内のブロックの順は、全話者において静音下条件を最初に実施し、その後の雑音下と残響下条件の順は各話者間でカウンターバランスをとった。各ブロックでは刺激語は擬似ランダムな順番で提示され、参加者によって各刺激語は10回ずつ発話された。

音声信号は、サンプリング周波数 48 kHz、量子化ビット数 32 bit にて、超単一指向性のコンデンサーショットガンマイク (RODE, NTG-2) を用いて録音した。マイク信号はマイクアンプ (PreSonus, DIGIMAX FS) およびオーディオインターフェース (RME, Fireface 802) を経由し、PC (OLIOSPEC, canarino fils5 Rev.3) 上でデジタル・オーディオ・ワークステーション (DAW; Adobe, Audition 2024) を用いて収録した。マイクは話者の口から約 25

cm の位置に設置し、全話者および全セッションで録音機器の設定を統一した。この手続きにより、可能な限り一貫した録音条件を維持した。よって録音音声から得られた平均インテンシティの値は参考値として扱うことが可能である。本研究では、声帯振動を測定するために electroglottography (EGG; Glottal Enterprises, EG2-PCX2) を使用した。この EGG データは、VOT の音響分析において補助的なデータとして用いた。なお雑音下条件においては、等価騒音レベル 75 dB (A) の白色雑音を開放型ヘッドホン (SONY, MDR-MV1) を用いて提示した。残響下条件ではリバーブレータ (Roland, RSS-303) にて話者の発話音声に実時間で残響が付加され、音圧レベル 75 dB (A) にて同ヘッドホンから提示された。付加した残響の残響時間は約 2.92 s であった。

2.4 音響分析

各試行の発話音声に対して、Praat にてアノテーションを行い、Python のライブラリである Parselmouth を用いて音響分析を行った。*fo* に関連する指標は全て、全ての話者データに基づき *z* スコア化した後、各話者内で standardization した [10]。まず、発話全体における音響指標として、刺激語全体での平均インテンシティ、話速、平均 *fo*、*fo* レンジとターゲット子音の後続母音の持続時間を算出した。また、閉鎖音の有声無声対立に関わる音響的特徴として VOT と後続母音の onset の *fo* を音響指標とした。VOT の分析では、/b/ にて pre-voicing が発生した発話を統計分析から除外し、それ以外の /b/ と /p/ の発話では、閉鎖の解放から後続母音に伴う声帯振動の開始時点までの時間を計測した。この開始時点は、音声波形上のゼロ交差とスペクトログラム上の F1 の立ち上がりを基準とし、判断が難しい場合は補助的に EGG データを参照し、総合的に判断した。また、後続母音の持続時間の 10% 時点までの平均 *fo* を算出し、これを後続母音の onset の *fo* とした [8]。

2.5 統計分析

統計分析には R (4.5.3) を用いた。各音響指標を従属変数とする線形混合モデルを *lme4* パッケージを使用して作成した。モデルは個別に適合した。独立変数は、発話環境、刺激語あるいはターゲット子音、発話スタイル、性

別, およびそれらの交互作用を含めた。そして話者をランダム効果に設定した。固定効果の検定には *lmerTest* パッケージ, 事後検定には *emmeans* パッケージを用いた。 p 値は Tukey 法にて補正した。VOT と後続母音の持続時間の分析には話速を共変量とした。

3 結果

3.1 発話全体における音響指標の変化

平均インテンシティは, Clear speech の主効果が全ての刺激語で有意であった (全て $p < 0.001$)。発話環境における主効果は, 雑音による主効果が有意であり ($p < 0.001$), 雑音と Clear speech の交互作用は /b/ の HL のみ有意であった ($p < 0.01$)。残響の主効果は /p/ の HL のみ有意であり ($p < 0.05$), Clear speech と残響の交互作用は有意でなかった。

話速は, Clear speech の主効果は /m/ と /p/ で有意であった (/p/-LH: $p < 0.05$, /p/-HL: $p < 0.01$, /m/-LH: $p < 0.01$, /m/-HL: $p < 0.001$)。発話環境は, 雑音の主効果は /p/ と /m/ を語頭に含む HL の刺激語にて生じ (/p/: $p < 0.001$, /m/: $p < 0.01$), 雑音と Clear speech の交互作用は /b/ を語頭に含む HL の刺激語にて生じた ($p < 0.05$)。残響の主効果は全ての刺激語で生じた (/b/-HL: $p < 0.01$, その他全て $p < 0.001$)。残響と Clear speech の交互作用は生じなかった。

後続母音の持続時間は, Clear speech の主効果は有意であり ($p < 0.05$), 発話環境は有意ではなかった。しかし, 残響と Clear speech における交互作用は有意であった ($p < 0.05$)。

平均 fo は, Clear speech の主効果は有意であった ($p < 0.01$)。また fo レンジについても Clear speech の主効果は有意であった ($p < 0.001$)。発話環境は雑音の主効果が有意であり ($p < 0.05$), 雑音と Clear speech による交互作用が有意であった ($p < 0.05$)。

3.2 有声無声対立における音響指標

各刺激語での VOT の変化を Fig. 1 に示す。LMM の結果は VOT では, ターゲット子音による主効果は有意であった ($\beta = 20.91$, s.e. = 2.80, $t = 7.47$, $p < 0.001$)。Clear speech による主効果も有意であった ($\beta = -10.62$, s.e. = 2.75, $t = -3.86$, $p < 0.001$)。発話環境の主効果は雑音下にて有意であった ($\beta = -10.24$, s.e. = 2.87, $t = -3.57$, $p < 0.001$)。Clear speech と発話環境の交互作用は有意ではなかった。

一方, LMM の事後検定の結果から, /b/ から /p/ の値を減じた VOT の差では, 複数の条件間で有意な差が認められた。まず女性話者では LH にて, 通常発話における静音下と残響下の間で有意差があり ($\beta = 10.29$, s.e. = 3.81, $t = 2.70$, $p < 0.01$), 雑音下と残響下の間でも有意差があった ($\beta = 9.51$, s.e. = 3.67, $t = 2.59$, $p < 0.01$)。Clear speech では LH にて, 静音下と残響下の間で有意であった ($\beta = 8.69$, s.e. = 3.63, $t = 2.39$, $p < 0.05$)。次に男性話者では LH にて, 通常発話において静音下と雑音下の間で有意差があった ($\beta = -9.23$, s.e. = 4.45, $t = -2.07$, $p < 0.05$)。

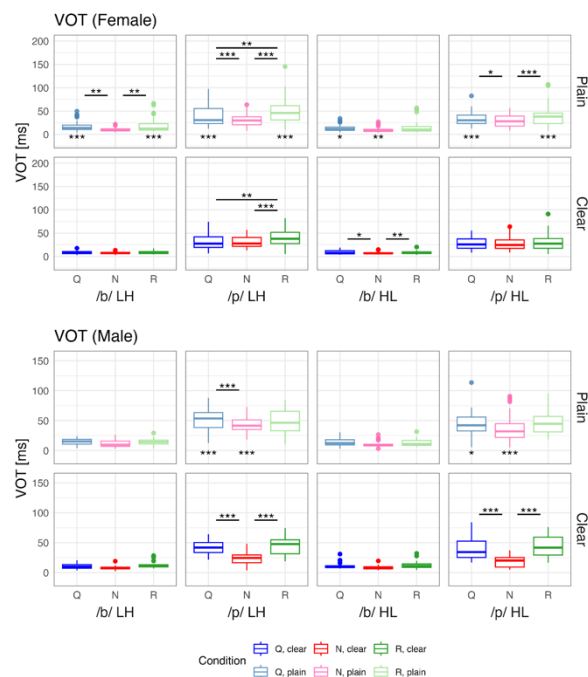


Fig. 1 Effects of speaking environment on VOT.

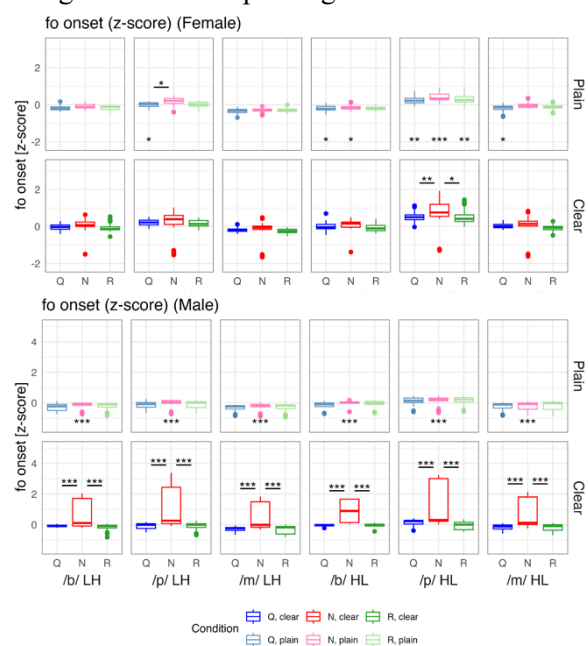


Fig. 2 Effects of speaking environment on following-vowel onset fo .

0.05)。Clear speech では LH にて、静音下と雑音下の間で有意差があり ($\beta = -11.17$, s.e. = 4.45, $t = -2.51$, $p < 0.05$), 雑音下と残響下の間でも有意差があった ($\beta = 15.28$, s.e. = 4.26, $t = 3.59$, $p < 0.001$)。HL では、Clear speech のみ静音下と雑音下の間で有意差があり ($\beta = -14.52$, s.e. = 5.41, $t = -2.69$, $p < 0.01$), 雑音下と残響下の間で有意差があった ($\beta = 17.96$, s.e. = 5.76, $t = 3.12$, $p < 0.01$)。

各刺激語での後続母音の onset の f_0 の変化を Fig. 2 に示す。後続母音の onset の f_0 では、刺激語の主効果は有意であったが ($\beta = 0.32$, s.e. = 0.09, $t = 3.59$, $p < 0.001$), Clear speech による主効果は有意傾向であった ($\beta = 0.15$, s.e. = 0.08, $t = 1.79$, $p = 0.074$)。交互作用は Clear speech と性別 (男性) と雑音にて有意であった ($\beta = 0.85$, s.e. = 0.17, $t = 4.85$, $p < 0.001$)。またターゲット音節/p/・Clear speech・雑音・HLの間でも交互作用が有意であった ($\beta = 0.49$, s.e. = 0.24, $t = 2.05$, $p < 0.05$)。一方、LMM の事後検定の結果、/b/から/p/の値を減じた後続母音の onset の f_0 の差は、男性話者の LH にて Clear speech では静音下と雑音下で有意差が生じた ($\beta = 0.30$, s.e. = 0.15, $t = 2.01$, $p < 0.05$)。

4 考察

発話全体における音響指標では、多くの指標で Clear speech による有意な差があり、先行研究と同様の発話変化 [1]が観察された。一方、発話環境と Clear speech の交互作用は一貫して生じず、発話スタイルを統制した際の発話環境による付加的な変化は限定的であり、先行研究と同様の傾向であった [6]。

語頭閉鎖音の有声無声対立に関わる音響指標においても、まず VOT については Clear speech の効果が有意であったが、発話環境と Clear speech の交互作用に有意差は生じず、発話環境が Clear speech を増強するような傾向は顕著ではなかった。しかし、/p/と/b/間の対立に着目すると、特に雑音下と残響下の間で有意な差があった。この傾向について、Fig.1 に示すように、特に/p/の VOT は静音下と比較して雑音下では短くなる一方、残響下では長くなる傾向を示した。このように雑音下と残響下で静音下と比較した際に反対の傾向を示すことが、VOT の対立における差を生んだ可能性がある。また本研究の雑音下における

/p/の VOT 短縮傾向は、オランダ語母語話者と英語母語話者で観察された無声閉鎖音での Lombard speech の傾向と類似している [11]。

後続母音の onset の f_0 は、まず/p/と/b/間で有意な差があり、/p/の方が高くなる傾向があった。これは先行研究 [8]と同様の傾向であった。Clear speech の主効果が有意ではなかった理由は、男女の傾向の差、 f_0 変化傾向が LH と HL 間で必ずしも一致しないことや、発話環境間で Clear speech の効果が異なることが反映された可能性がある。交互作用の傾向は一貫しなかった。/p/で生じた交互作用から、/p/では Clear speech、雑音下と HL の組み合わせにて f_0 が増加する傾向が示されたが、この変化は/b/-/p/の対立拡大に結び付かなかった。

これらの結果から、発話全体における音響指標の変化は先行研究 [6]と同様、発話スタイルを統制すると発話環境による影響は限定的であった。有声無声対立に関わる音響指標では、/p/-/b/間の差に注目すると VOT では、話者の性別や発話スタイルで変化の程度の違いはあるが、全体の傾向として発話環境によって変化する方向が異なることが示唆された。

謝辞

本研究は上智大学学術研究特別推進費 (重点領域研究) の助成を受けた。また本研究は上智大学「人を対象とする研究」に関する倫理委員会の承認を受けて実施した (2025-017)。

参考文献

- [1] Smiljanić and Bradlow, *Lang. Linguist. Compass.*, 3 (1), 236–264, 2009.
- [2] Junqua, *Speech Commun.*, 20(1–2), 13–22, 1996.
- [3] 程島他, IEICE Technical Report, SP2009-69, 43–48, 2009.
- [4] 久保, 赤木, IEICE Technical Report, SP2017-91, 39–44, 2017.
- [5] 森田, Thesis or Dissertation, 北陸先端科学技術大学院大学, 1–34, 2020.
- [6] Ishikawa *et al.*, *Sci. Rep.*, 14 (1), 20069, 2024.
- [7] 川井他, 音講論 (秋), 1077–1078, 2025.
- [8] Gao and Arai, *J. Phon.*, 77, 100932, 2019.
- [9] 高田, 『日本語の語頭閉鎖音の研究 — VOT の共時的分布と通時的変化』, くろしお出版, 2011.
- [10] Pinget and Quené, *J. Acoust. Soc. Am.*, 154(4), 2124–2136, 2023.
- [11] Marcoux and Ernestus, *J. Phon.*, 102, 101281, 2024.