

B-3

アクセントの誤りがもたらす語の認知の障害：意味プライミングによる検討

有賀照道

東京大学 総合文化研究科
ariga.terumichi@gmail.com

要旨

日本語の音声単語認知においてアクセントは重要な役割を果たす。逆に、アクセントの誤りは適切な単語認知を障害する要因となりうる。本研究では、アクセントの誤りが語の認知にどのような影響を与えるかを、「たんす HLL」のようにアクセントの誤り（「たんす LHH」の誤り）と音素の誤り（「ダンス HLL」の誤り）の2通りの解釈が存在する音声の認知の過程を感覚交差意味プライミング課題により探索することによって検討した。「たんす HLL」という音声刺激の後に「たんす LHH」に関連する「家具」を提示した場合、語彙性判断にかかる反応時間は大きく増大したが、「ダンス HLL」に関連する「音楽」を提示した場合はそれほど大きな増大は見られなかった。このことから、アクセント情報と音素情報が矛盾する音声提示された場合、音素情報が正しくアクセントの誤りと見なした場合は語の認知が大きく障害されるが、アクセントが正しく音素情報の誤りと見做す場合はそれほど大きな障害とならない可能性が示唆された。

1. 背景

日本語の音声単語認知において、アクセントは目的の語へのアクセスを達成するための重要な役割を果たす。アクセントによって弁別される同音異アクセント語（e.g., 「じどう HLL」（児童）と「じどう LHH」（自動））は、アクセント情報に基づいて弁別的に認知される（Ariga, 2022 ; Sekiguchi & Nakajima, 1999）。音声の一部の情報が得られない場合にも、アクセント情報に基づいて単語全体を予測的に認知することが可能である（Cutler & Otake, 1999）。

このことは逆に、アクセントの誤りは、適切な単語認知を障害する要因となりうることを意味する。アクセントを誤った発話は、アクセントが正しい発話に比べて正しく認識できる割合が低下する（Minematsu & Hirose, 1995）。複数の脳波研究においても、単語に対するアクセント情報の誤りに対応する事象関連電位が報告されている（Hirose et al., 2021 ; Koso & Hagiwara, 2009 ; Koso et al., 2011）。一方、聞き手はアクセントが誤っていた場合でも語の認知が可能である。聞き手自身の知識と異なるアクセントで語が提示された際に、アクセント情報によらず音素情報のみに注目して語を認識する能力は、発達の早い時期から獲得される（山本 2012 ; Yamamoto & Haryu, 2018）。さらに、文提示下において文脈情報が誤りを修正しうる場合は、意味違反を起こすことなくアクセントを誤った単語が理解される（Tamaoka et al., 2014）。方言などを理由とした話者間のアクセントの知識に関する違いは、コミュニケーションを著しく困難にする要因とはなりえない（柴田 1961）。すなわち、アクセントの誤りは語の認識に影響を与えるものの、日本語の音声単語認知にはその誤りを適切に把握して適切な語を認知するメカニズムが備わっていると考えられるが、この誤ったアクセントで語が提示された場合の認知の過程はあまり検討されていない。

日本語において、アクセントの誤りは2通りに分類することが可能である。一つは、アクセントの誤りが別の有意味語となる場合である。「じどう HLL」（児童）という語を「じどう LHH」と発音した場合、それは「児童」を意味する語としては誤っているものの「自動」という別の単語としては適切な発音となる（Ariga, 2022 ; Sekiguchi & Nakajima, 1999）。もう一つは、アクセントの誤りが無意味語を作り出す場合である。「たんす LHH」という語を「たんす HLL」と発音した場合、それは単に誤った発音と

してしか解釈の余地がなくなる。

後者の場合において、アクセントを誤ったことにより生じた無意味語を聞き手がどのように理解するのか、先行研究では明らかになっていない。Ariga & Matsubara (2023) は、「たんす HLL」のようなアクセントを誤った発話において、語頭子音の VOT を操作して「たんす HLL」から「ダンス HLL」へ連続的に変化する音声を作成し、語頭子音の有声性を判断する課題を行った。その結果、語頭子音はアクセント情報を考慮した場合に語が有意味となるように解釈される傾向にあり (Ganong, 1980)、音響的に曖昧となった音素情報よりもアクセント情報に基づいて語が解釈されることが示唆された。しかし、この先行実験においても、VOT が曖昧でない場合に語頭子音の有声性の判断を行わせた場合、アクセント情報によらずに語頭子音の有声性を正しく判断することができた。すなわち、音素を知覚する課題においては「たんす HLL」という発音を聴取した場合に、アクセント情報に影響されることなく語頭の音素を /t/ と認識することが可能であった。

しかし、この結果は音素に音響的曖昧性のない「たんす HLL」という音声は「たんす LHH」(箆笥)を意図した発話として認識されたことを必ずしも意味するものではない。「たんす HLL」という発話は、アクセント情報と音素情報(語頭子音の有声性)が矛盾した発話であり、アクセントの誤りと音素の誤りの 2 つの誤りが交絡した発話となっている。このとき、「たんす HLL」という発音を誤った発話をあえて正しく解釈しようとした場合に 2 通りの方法が考えられることになる。一つは、この発話をアクセントの誤りと判断して、アクセント情報を修正して「たんす LHH」(箆笥)として認識するものである。もう一つは、この発話をアクセントが正しく、むしろ語頭の音素を誤ったものと判断して、音素情報を修正して「ダンス HLL」として認識するものである。

本研究では、「たんす HLL」のようにアクセントの誤りと音素の誤りの 2 通りの解釈が交絡する単語発話がどのように解釈されるかを、感覚交差意味プライミング実験によって検討する。本研究は 2 つの実験からなる。実験 1 では、「たんす HLL」のような発話がアクセントの誤り(「たんす LHH」の誤り)として処理される可能性を検討する。実験 2 では、同じ発話が音素の誤り(「ダンス HLL」の誤り)として処理される可能性を検討する。

2. 実験

2.1 参加者

実験参加者は実験バイト (<https://www.jikken-baito.com/>) を用いて募集した。実験 1 は東京方言母語話者 25 名(平均年齢 23.36 歳 ($SD = 4.39$)), 実験 2 は東京方言母語話者 24 名(平均年齢 25.71 歳 ($SD = 4.81$)) が参加した。

2.2 実験刺激

「たんす LHH」/「ダンス HLL」のように、アクセント型と語頭子音の有声性の対立のある 3 モーラ語を 12 セット用意した。これらのアイテムでは、アクセント型と語頭子音の有声性は語彙的に連動する。すなわち、「たんす LHH」/「ダンス HLL」は有意味語であるが、「たんす HLL」/「ダンス LHH」は無意味語となる。これらの刺激語に対する語彙活性を意味プライミングによって検討するため、刺激語ペアの片方に意味的に関連し、もう一方に意味的に関連しない漢字 2 字の単語 (e.g., 「たんす LHH」に対して「家具」, 「ダンス HLL」に対して「音楽」) を用意し、さらに、意味プライミング効果を測定するための統制条件に用いる語として、この 2 語のどちらとも関連しない語 (e.g., 「さばく」(砂漠))

を用意した。

用意した刺激語ペアの意味的関連度の妥当性を、実験に参加しない 11 名の日本語母語話者による評定実験によって確認した。参加者は刺激語ペアの意味的関連度を 1 から 6 (6 が最も関連度が高い) で評価した。評定実験の結果、意味的関連のある条件の z 化された平均評定値は 1.20, 関連のない条件と統制条件の平均評定値はそれぞれ -0.54 , -0.64 であった。LME による統計分析の結果、関連あり条件の評定値は統制条件の評定値よりも有意に高い一方 ($\beta = 1.847$, $SE = 0.075$, $df = 11.252$, $t = 24.776$, $p < .001$), 関連なし条件と統制条件の間には評定値に有意な差は見られなかった ($\beta = 0.101$, $SE = 0.118$, $df = 10.936$, $t = 0.857$, $p = .410$)。

この刺激語をもとに、表 1, 2 に示した実験条件を作成した。congruent 条件は、プライムの単語が正しく発音されており、その単語に意味的に関連する単語がターゲットとして提示される (e.g., 「たんす LHH 一家具」「ダンスー音楽」)。mispronounced 条件では、誤った発音の単語がプライムとして提示される。実験 1 では、congruent 条件の単語のアクセント型を置き換えたアクセントの誤りの語 (e.g., 「たんす LHH」を「たんす HLL」, 「ダンス HLL」を「ダンス LHH」), 実験 2 では、語頭の子音の有声性を置き換えた音素の誤りの語 (e.g., 「たんす LHH」を「ダンス LHH」, 「ダンス HLL」を「たんす HLL」) をそれぞれプライムとして提示した。baseline 条件では、意味プライミング効果を測るための統制群として、ターゲットの語に意味的に関連のない語 (「さばく」(砂漠)) をプライムとして提示した。アクセント型の効果を確認するため、正しい発音である congruent 条件の語のアクセント型が平板型 (unaccented) であるものと頭高型 (accented) であるものの 2 条件を用意した。

1 アイテム 6 条件のうち刺激語の重複しない 2 条件を選び、1 名の参加者につき 24 試行をテスト試行とした。72 試行のフィラー試行 (このうち、ターゲットが無意味語であるものは 48 試行) を追加し、全体で 96 試行をランダムな順序で提示した。

プライムの音声刺激は、発音を誤った単語が用いられる mispronounced 条件を含め、音韻論の知識の

表 1 実験刺激 (実験 1: アクセントの誤り)

Prime Pattern	Accent Type			
	unaccented		accented	
	Prime	Target	Prime	Target
congruent	たんす LHH		ダンス HLL	
mispronounced	たんす HLL	家具	ダンス LHH	音楽
baseline	さばく		さばく	

表 2 実験刺激 (実験 2: 音素の誤り)

Prime Pattern	Accent Type			
	unaccented		accented	
	Prime	Target	Prime	Target
congruent	たんす LHH		ダンス HLL	
mispronounced	ダンス LHH	家具	たんす HLL	音楽
baseline	さばく		さばく	

ある東京方言母語話者(女性)による自然発話を用いた。ただし *mispronounced* 条件においては, *congruent* 条件の単語に対してアクセントを入れ替えて発話する (i.e., アクセントの誤りとして発話する) ように教示し, 録音後に実験者が確認して聴覚印象として正しく別のアクセント型で語が発話されていることを確認した。*mispronounced* 条件を含む全条件の音声刺激は, 実験 1 と実験 2 で同一の録音を用いた。刺激は防音室において, Samson Q2U マイクを使用し, Praat 6.2.14 (Boersma & Weenink 2022) を用いて量子化ビット 16 bit, 標準化周波数 44.1 kHz で録音された。録音後に音声のインテンシティを標準化し, ファイルを mp3 形式で保存した。

2.3 手順

実験は PCibex (Zehr & Schwarz, 2018) を用いて, オンラインで実施した。各試行において, 注視点 (+) を 1 秒間提示した後, プライムの音声刺激を提示した。プライムのオフセットで, 注視点を画面から消し, ターゲットの視覚刺激を提示した。実験参加者はターゲットの刺激語の語彙性判断をできるだけ速く行うように教示された。実験参加者が回答を行うか, 3 秒間回答がないまま経過した場合にターゲットの視覚刺激が画面から消え試行が終了した。

語彙性判断の終了後, プライムの音声刺激へ注意を向けさせるために, ダミーの音声判断課題を行った。この音声判断課題では, 直前に聞いたプライムの音声刺激が 3 モーラであったか 4 モーラであったかを判断させた。なお, 実験開始時の教示において, モーラ概念について簡潔に説明した。音声判断課題の終了後, 1 秒の空白を開けて次の試行を開始した。

語彙性判断課題の後, 実験参加者のアクセントの知識を確認するため, 発音課題を実施した。実験で使用した単語を画面上に提示し, 発音させた。実験の所要時間は全体で約 30 分であった。

2.4 結果

4 名の参加者 (実験 1 : 2 名, 実験 2 : 2 名) のデータは, 語彙性判断課題における正答率が著しく低かった (60%未満) ため分析から除外した。また, 2 つのアイテムのデータは発音課題においてアクセントを誤って回答した参加者が多くいたため, 分析の対象外とした。参加者が発音課題において発音を誤った単語をプライムとして用いた試行, 反応時間が 300 ms 未満または 1300 ms 以上である試行のデータも除外した。

語彙性判断課題における反応時間を LME により分析した。プライム音声のアクセント型 (Accent Type) にはエフェクトコーディングを用い, *unaccented* 条件を -0.5 , *accented* 条件を 0.5 とコード化した。プライムの種類 (Prime Pattern) にはダミーコーディングを用い, *baseline* 条件と *congruent* 条件の比較, *baseline* 条件と *mispronounced* 条件の比較についてそれぞれ 0, 1 とコード化した。最適モデルを選択したのち, モデル推定値と観測値との間の残差が $2.5 SD$ を超えるデータをさらに除外した。

各実験における条件ごとの平均反応時間を図 1 に示した。さらに, 実験ごとの LME による反応時間の分析の結果を表 3, 4 にそれぞれ示した。実験 1 では, *mispronounced* 条件の反応時間は *baseline* 条件よりも有意に長かった ($p = .047$)。しかし, *congruent* 条件と *baseline* 条件間に反応時間の差はなく ($p = .301$), アクセント型の主効果 ($p = .474$), プライムの種類とアクセント型の交互作用 (*congruent* × *Accent Type* : $p = .322$; *mispronounced* × *Accent Type* : $p = .158$) は有意ではなかった。アクセントの発音を誤った実験 1 の *mispronounced* 条件では, *baseline* 条件よりも語彙性判断にかかる時間が増加した。これはアクセントの誤り (e.g., 「たんす LHH」を「たんす HLL」と発音する) が語彙活性を大きく阻害する要因

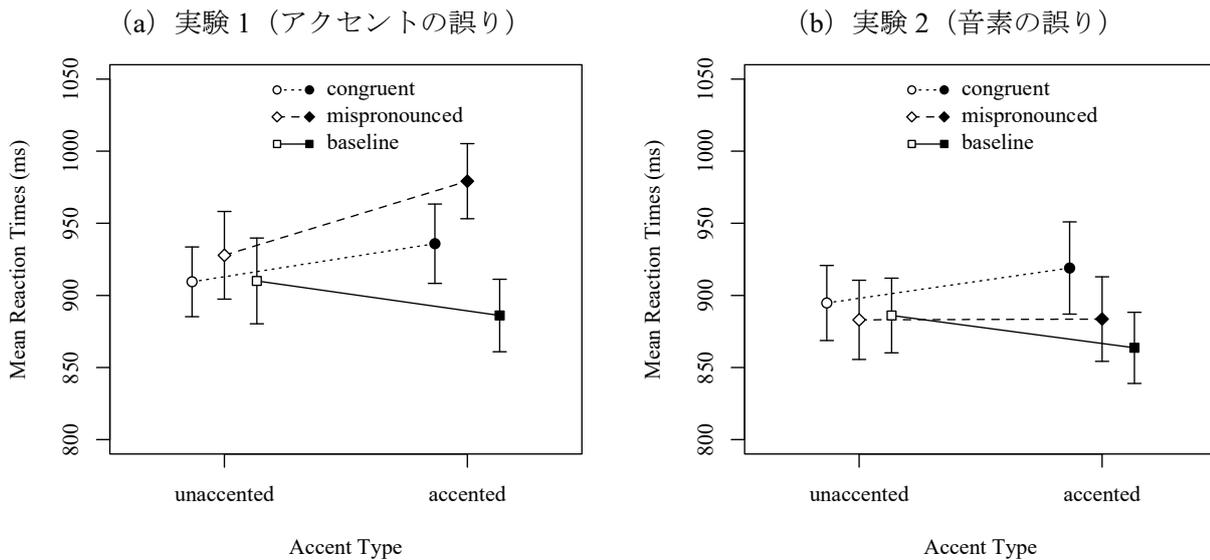


図1 各実験 (a : 実験 1, b : 実験 2) における条件ごとの平均反応時間 (ms)。エラーバーは標準誤差 (SE) を示す。

表3 LME による検定結果 (実験 1)

	β	SE	df	t	p	
(Intercept)	913.587	25.530	33.425	35.785	< .001	***
congruent	25.879	24.956	261.338	1.037	.301	
mispronounced	50.157	25.172	251.992	1.993	.047	*
Accent	-24.596	34.296	256.088	-0.717	.474	
congruent \times Accent	48.568	48.926	252.026	0.993	.322	
mispronounced \times Accent	71.562	50.540	255.676	1.416	.158	

Significant codes: *** $p < 0.001$, * $p < 0.05$

表4 LME による検定結果 (実験 2)

	β	SE	df	t	p	
(Intercept)	901.650	35.821	31.368	25.171	< .001	***
congruent	33.197	19.024	286.976	1.745	.082	†
mispronounced	12.757	19.251	287.162	0.663	.508	
Accent	-15.411	25.583	286.003	-0.602	.547	
congruent \times Accent	28.916	38.004	286.790	0.761	.447	
mispronounced \times Accent	37.356	38.455	286.876	0.971	.332	

Significant codes: *** $p < 0.001$, † $.05 < p < .1$

となることを示唆するものである。

一方, 実験 2 では, mispronounced 条件の反応時間は baseline 条件と差がなく ($p = .508$), むしろ congruent 条件と baseline 条件の反応時間の差が有意傾向であった ($p = .082$)。実験 1 と同様に, アクセント型の

主効果 ($p = .547$), プライムの種類とアクセント型の交互作用 ($\text{congruent} \times \text{Accent Type} : p = .447$; $\text{mispronounced} \times \text{Accent Type} : p = .332$) は有意ではなかった。音素 (語頭子音の有声性) の発音を誤った実験 2 の *mispronounced* 条件では, 実験 1 と異なって *baseline* 条件と比較した場合の語彙性判断にかかる時間の増加が見られなかった。*mispronounced* 条件は *baseline* 条件に比べて反応時間が減少しているわけではないことから, 発音の誤り自体は語彙活性に影響を与えていると考えられる。しかしながら, アクセントの誤り (実験 1) と比較すれば, 音素の誤り (実験 2) はより軽微な誤りとして処理されることが示唆される。

3. 考察

「たんす HLL」のようにアクセント情報と音素情報が矛盾した発話を聴取した場合, それをアクセントの誤りとして解釈した場合の意味的関連語をターゲットとして提示した際 (e.g., 「たんす HLL 一家具」: 実験 1) の語彙性判断は統制条件と比べて大きく阻害されたのに対し, 音素の誤りとして解釈した場合の意味的関連語 (「たんす HLL 一音楽」: 実験 2) の語彙性判断にはそれほど大きな阻害は見られなかった。このことから, 「たんす HLL」のようなアクセント情報と音素情報が矛盾する発話が, 音素情報が正しくアクセントが誤っている (「たんす LHH」の誤り) よりも, アクセント情報が正しく音素情報が誤っている (「ダンス HLL」の誤り) ものとして解釈される傾向にあることが示唆される。

ただし, 本実験の結果に関してはアクセント情報と音素情報に誤りのない *congruent* 条件のふるまいについて議論の余地が残る。*congruent* 条件では, 発音に誤りのない語が提示され, かつその語に意味的に関連する語が提示されているため, 通常の意味プライミング課題の枠組みにおいては, 意味プライミング効果により統制条件に比べて反応時間が短くなることが予想された (Ariga, 2022; Cutler, 1986)。しかしながら, 本実験では統制条件と反応時間に差がないか (実験 1), むしろ反応時間が統制条件よりも長くなる傾向が観察された (実験 2)。

この *congruent* 条件における意味プライミング効果の欠如には, 実験内において単純語の発話としてはありえない誤った語が出現することによる影響の可能性がある。*mispronounced* 条件で提示された語 (「たんす HLL」や「ダンス LHH」) は, 東京方言の単純語としては, アクセントまたは音素の誤りを含む不適切な発話である。しかし, これらの発話は東京方言であっても複合語の一部の要素としては出現する可能性がある (e.g., 「ダンス-パーティーLHH-HLLL」)。提示された音声は複合語の部分要素である可能性も考えた場合, アクセントや音素に誤りのない *congruent* 条件の発話に対しても, 単純語としての解釈と異なる意味で語を理解しなければならない場合が生じる。例えば, 「ようふく-だんす LHH-HLL」 (洋服箆笥) という複合語が提示された場合, 単純語としては本来「ダンス」と解釈されるべき後部要素の発話を「箆笥」として解釈することになる。実験で提示された音声刺激は全て単純語であったものの, 実験参加者は複合語の部分要素という解釈の可能性を考慮したため, 「ダンス HLL」という発話を聞いた場合に, 「ダンス」と「箆笥」の両方に語彙アクセスを行わなければならないと, 結果的に意味プライミング効果が無効化された可能性がある。

しかしながら, 実験 1 と実験 2 における *mispronounced* 条件のふるまいの変化は, アクセント情報と音素情報が対立する場合に, アクセント情報に基づいて語の解釈を行うほうが単語認知において有利にはたらくことを示唆する。音声単語認知は, 言語情報の冗長性によりさまざまな妨害に対して頑健なシステムとなっている。言語情報の冗長性に注目したとき, 単語中で時間的に一部分を占めるに過ぎない音素情報は容易に失われうるのに対して, 単語全体にわたって提示されるアクセント情報は喪失に対し

て頑健である。この考えに基づけば、音素とアクセントの矛盾が存在した場合に、容易に失われる音素情報に注目するよりも、失われにくいアクセント情報に注目する方が安定的に語を解釈することが可能であると考えられる。したがってアクセントの誤りは語アクセスを大きく阻害する要因となりうるが、アクセントが正しいとした上で音素が誤ることはそれほど大きな影響を与えないと考えられる。このことは、日本語の音声単語認知においてアクセントが適切に提示されることの重要性を示すものである。

参考文献

- Ariga, T. (2022). Pitch accent constrains lexical activation in Japanese spoken word recognition: A semantic priming study. *Language and Information Sciences*, 20, 1–17.
- Ariga, T., & Matsubara, R. (2023). Top-down effects of lexical pitch accent on phonetic categorization in Japanese. In: R. Skarnitzl, & Jan Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 210–214). Guarant International.
- Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer. <https://www.fon.hum.uva.nl/praat/>
- Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, 29(3), 201–220.
- Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *The Journal of the Acoustical Society of America*, 105(3), 1877–1888.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6(1), 110–125.
- Hirose, Y., Kobayashi, Y., Chen, T., Ito, A., & Ito, T. (2021). ERP responses to different types of pitch accent violation in Tokyo Japanese: Rule application or lexical memory? In H. Jeon (Ed.), *Japanese/Korean linguistics vol. 28* (pp. 333–344). CSLI Publications.
- Koso, A., & Hagiwara, H. (2009). Event-related potential evidence of processing lexical pitch-accent in auditory Japanese sentences. *Neuroreport*, 20(14), 1270–1274.
- Koso, A., Ojima, S., & Hagiwara, H. (2011). An event-related potential investigation of lexical pitch-accent processing in auditory Japanese. *Brain Research*, 1385, 217–228.
- Minematsu, N., & Hirose, K. (1995). Role of prosodic features in the human process of perceiving spoken words and sentences in Japanese. *Journal of the Acoustical Society of Japan (E)*, 16(5), 311–320.
- Sekiguchi, T., & Nakajima, Y. (1999). The use of lexical prosody for lexical access of the Japanese language. *Journal of Psycholinguistic Research*, 28(4), 439–454.
- 柴田武 (1961) 「日本語のアクセント」『言語生活』117, 14–20.
- Tamaoka, K., Saito, N., Kiyama, S., Timmer, K., & Verdonschot, R. G. (2014). Is pitch accent necessary for comprehension by native Japanese speakers? –An ERP investigation. *Journal of Neurolinguistics*, 27(1), 31–40.
- 山本寿子 (2012) 「誤ったアクセントで発音された単語に対する幼児の認知」『教育心理学研究』60(2), 127–136.
- Yamamoto, H. W., & Haryu, E. (2018). The role of pitch pattern in Japanese 24-month-olds' word recognition. *Journal of Memory and Language*, 99, 90–98.
- Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). <https://www.pcibex.net/>