

研究ノート

単語境界明示化に關与する音響的手がかりについて（文献調査）

On the acoustic cues for the word segmentation process: A literature review

山本 勝巳

要約：音声言語には書き言葉にはない様々な特徴がある。本論文ではそのうち言語リズムや高さ、ストレスといった韻律的特徴について述べ、それらがリスニング過程においてどのような役割を担っているのかについて先行研究のレビューを行なう。

Key Words：言語リズム・リスニング・gating task

1. はじめに

現在、筆者は日本人英語学習者が音声として提示された英語発話から単語を切り出すための手がかりとする言語的特徴に興味を持ち、実験的検討を進めている。そこでそれに関連する先行研究のうち、言語リズム・リスニングのプロセスに関するものをまとめた上で、音声信号から単語列を切り出すプロセスに関する検討の必要性について論じることとした。

2. 言語リズム

言語のリズムに関して言及した初期の報告には、例えば Pike (1945) がある。そこでは stress-timed な言語においては、ある stress から次の stress までがほぼ等間隔になるように並び、その一方で syllable-timed な言語では syllable がほぼ等間隔に並ぶ傾向がある、と述べられている。それを受けて Abercrombie (1967, p. 96) には "As far as is known, every language in the world is spoken with one kind of rhythm [a syllable-timed rhythm] or with the other [stress-timed rhythm]." と記述されている。

しかし、音声言語の音響的測定データの面からこうした記述に対応する分析結果を提出することは困難なことであった (stress-timed な言語の例とされる英語に関する音響分析の歴史については、例えば Lehiste (1977) を参照のこと)。Mora-timed な言語として言及される日

本語についても、mora (モーラ) に相当する音声刺激の長さを測定した Beckman (1982) に示されるように、モーラ自体の等時性には疑問が多い。むしろ、Homma (1981) や Port *et al.* (1987) に示されるように、単語のモーラ数を増やしていくとほぼ一定の割合で単語長が増加したり、同じモーラ数の単語の平均長が同程度に落ち着くといった報告が多く、実験データから単独モーラの等時性を示すことは困難である。その後、Dauer (1983) によって異なる言語リズムに属するとされてきた複数の言語の実測データが示され、言語リズムの存在を支持する単純な物理的根拠は存在しないというのが現在の基本的な考え方だと思われる。一方で、測定の方法を変えることで言語リズムを支持する何らかの音響的実態が存在するとする報告はその後もいくつか提出されている。これには例えば、Port *et al.* (1996) や Grabe & Low (2002) などがある。

これに対して、異なる言語リズムの言語話者の発話・知覚様式から、言語リズムの心理的実在にアプローチする研究の流れも存在する。その中で大きな影響を持っているのが Cutler らによる一連の報告である (例えば Cutler & Norris (1988) などを参照のこと)。彼女らは英語・フランス語・日本語などの異なる言語リズムを持つ言語話者がそれぞれの韻律単位に基づいて話しことばの単語境界を分節しているとの報告を数多く挙げている。言いかえると、彼女らの関心は分節単位 (segmentation unit) の追求にあると考えられる。単語境界明示化に対する彼女らの心理言語学からのアプローチに対して、音響分析に基づく報告には例えば Turk &

Shattuck-Hufnagel (2000) などがあり、そこでは分節音の長さによる手がかりの存在が指摘されている。

3. リスニングのプロセス

言語情報を処理するには心的辞書に蓄積されている語彙情報にアクセスする必要があると考えられる。Levelt (1993) は母語話者の言語処理過程について心的辞書(レキシコン)に着目した次のような図を提示している(図1)。

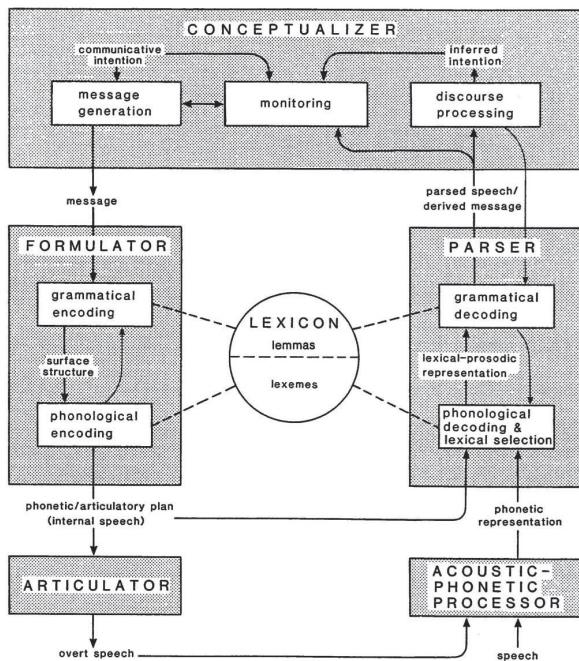


図1. 言語処理過程の模式図 (Levelt, 1993, p. 2より)

この図を踏まえて Kadota (2003, pp. 24-25) にはリスニングにおける各部門の働きについて次のような説明が見られる(詳細については Yamamoto (2006) を参照)。

(1) 音響・音声処理装置により、音声信号を音声表象に変換する。

(2) 文解析装置内の音韻化・語彙選択機構で、メンタルレキシコンから語彙項目内の lexeme 部から音韻情報が検索されて、音声表象を、一定の単位に分節し、語という単位を抽出することが可能になる。語彙選択において、活性化された複数の語候補から絞り込みを行い、最終的に特定の語彙項目を取り出すという語彙アクセスが行われる。音韻化・語彙選択機構のアウトプットは、語彙が挿入された、語彙・韻律表象である。

(3) 語彙・韻律表象は次なる高次の文法的コード化のインプットとなる。

(4) 文解析装置の結果生み出される文の概念表象から話し手の発話意図の最終的な解釈にむけて、概念化装置内の談話処理機構が作動する。また必要に応じて、発話者の発話意図の推測が行われる。

Yoneyama (2002, p. 23) は“lexical access”に「語彙接近」という訳をあてた上で「音声情報と心的辞書(Lexicon)に蓄積されている語彙情報との照合が行われる過程」と定義しているが、これは Kadota (2003) の言う(2)の段階であろう。

こうした言語処理モデルを踏まえた日本人英語学習者の言語処理過程に対する先行研究は読みに関するものが多いが、音声言語の処理に際しても同様のプロセスを経由する必要があると考えられる。Takano (1995, p. 255) が述べる「外国語効果」とは、「不慣れな外国語を使っている最中は、その外国語を使うのが難しいだけでなく、思考力も一時的に低下するという現象」であるが、こうした現象はワーキングメモリにおける複数タスクの同時実行の困難が原因であると考えられるからである。

ここで問題になるのは音声言語の場合、読みと異なり基本的に連続する音韻列から適切に単語を切り出す必要があるという点であり、この問題について次に論じる。

4. 音声信号から単語列を切り出すプロセスに関する検討

McQueen & Cutler (2001) には“spoken word access process”という処理段階への言及がある。これは「音声信号から単語列を切り出すプロセス (p. 470)」とされ、得られた単語列が「話しことばの基本的な知覚表象 (p. 470)」を形成し、「われわれが発話の意味を取るための解釈プロセスへの入力となる (p. 470)」と説明されている。さらに「語彙アクセスプロセスにおける基本的な要素の一つは単語の切り出しである (p. 480)」とした上で「個々の単語をどのようにとりだすのだろうか (p. 480)」との問題提起がある。Levelt (1993, p. 8) もこうした処理段階を明示してはいないものの「聞き手が連続音声処理する場合の最初の大きな問題はそれをどう区切るかである」と指摘している。

単語境界の知覚の手がかりについて、Yoneyama (2002) ではそれまでの Cutler らの研究から「言語の韻律構造を構成する韻律単位に基づく分節方法は、語彙境界の同定に有益 (p. 25)」と総括し、これを「音声の分節に関する言語普遍の法則 (p. 25)」と呼んでいる。あわせて、Cutler (1997) を引用して「成人話者は、語彙接近のために必要な妥当と思われる語彙境界を探すため

に可能な限りの音韻情報を用いていると想定されている (p. 27)」と述べている。

英語話者の単語分節に関して, Nakatani & Dukes (1977) では知覚面の手がかりは語頭にあることが示され, 続く Nakatani & Schaffer (1978) では 1. 単語を知っていること 2. phonotactics な情報 3. allophonic な情報 4. 強勢パターン・リズムが主要な手がかりとなっていることが示された。Cutler & Clifton, Jr. (2000) には「英語話者には韻律の示す手がかりよりも強勢に伴う音質といった分節音の手がかりが有効である (p. 137)」との指摘がある。一方, 中国語・日本語話者を用いた実験結果から, これらの話者では「語の活性化に際してトーンやピッチの情報が利用されている (p. 138)」との指摘もある。つまり単語の分節に対して機能する音響の特徴は母語によって異なる部分があると考えられるのである。例えば英語では stress という lexical prosody が語彙接近に用いられないとした Cutler らの報告に対して, Sekiguchi & Nakajima (1999) は日本語母語話者が lexical prosody であるピッチアクセントを語彙接近に用いているとの報告を行ない, 日本語と英語の lexical prosody の使われ方のちがいについて興味深い指摘を行っている。また Cutler & Otake (1999) も日本語母語話者を対象とした知覚実験から, 彼らの場合アクセント情報が語彙認識において候補語の活性化や選択に対する制限となっているとの指摘を行っている。

一方で, 分節単位の研究領域では例えば Norris *et al.* (2001, p. 638) には「リズム構造は母語の単語切り出しには役に立つが非母語の入力に対しては不適切な手がかりを与えることもある」との指摘がある。

分節単位の研究にも音響の手がかりについての研究にも利用される実験手法として gating paradigm の有効性が報告され (Grosjean (1980)), これを受けて gating task を手法として取り入れたいくつもの報告が行なわれている。Grosjean (1980) によればこの手法は “entails presenting a spoken language stimulus repeatedly and increasing its presentation time at each successive pass. (p. 267)” とされ, その適用範囲は “from a simple CV syllable to a complex sentence (p. 267)” までと幅広い。例えば先に述べた Cutler & Otake (1999) や Sekiguchi & Nakajima (1999) はこの手法を利用したものである。今後の筆者の実験でもこの手法は十分利用可能だと考えている。

先に述べたように, 日本語と英語はそれぞれの母語話

者が異なる韻律単位で単語切り出しを行っているという点で, 心理言語学的には異なる言語リズムに分類されてきた。Lexical prosody の観点からも日本語と英語はそれぞれ高さ, 強勢を音響の手がかりとする言語と報告されている。これらの先行研究から, 日本人英語学習者を対象として, 彼らが英語における単語境界をどのように知覚しているのかを実験的に検討することには十分意義があるものと考えられる。今後はまず日本人英語学習者が語強勢を知覚するための時間的推移 (time course) について検討した上で, 彼らが一定の長さの聴覚入力から音響の手がかりを利用してどのように単語境界を知覚しているのかについて, gating task を利用して分析を進めたいと考えている。

参考文献

- Abercrombie, D. (1967). *Elements of general phonetics*. Chicago, IL: Aldine Publishing Company.
- Beckman, M. (1982). Segment duration and the ‘mora’ in Japanese. *Phonetica: International Journal of Speech Science*, 39, 113-135.
- Cutler, A. (1997). The comparative perspective on spoken-language processing. *Speech Communication*, 21, 3-15.
- Cutler, A., & Clifton, C., Jr. (2000). Comprehending spoken language: a blueprint of the listener. In C. M. Brown, & P. Hagoort (Eds.), *The neurocognition of language* (pp. 123-166). New York: Oxford University Press.
- Cutler, A. & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113-121.
- Cutler, A. & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America*, 105(3), 1877-1888.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- Grabe, E. & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven (Ed.), *Laboratory phonology 7* (pp. 515-546). Mouton de Gruyter.
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, 28(4), 267-283.
- Homma, Y. (1981). Durational relationship between Japanese stops and vowels. *Journal of Phonetics*, 9, 273-281.
- Kadota, S. (2003). *Eigo no mental lexicon - Goi no kakutoku, shori, gakushu*. Tokyo: Shohakusha.

- Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics*, 5, 253-263.
- Levelt, W. J. M. (1993). The architecture of normal spoken language use. In G. Blanken, E. Dittman, H. Grimm, J. Marshall, & C. Wallesch (Eds.), *Linguistic disorders and pathologies: An international handbook* (pp. 1-15). Berlin: de Gruyter.
- McQueen, J. M. & Cutler, A. (2001). Spoken word access processes: An introduction. *language and Cognitive Processes*, 16(5), 469-490.
- Nakatani, L. H. & Dukes, K. D. (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America*, 62, 714-719.
- Nakatani, L. H. & Schaffer, J. A. (1978). Hearing "words" without words: Prosodic cues for word perception. *Journal of the Acoustical Society of America*, 63(1), 234-245.
- Norris, D., McQueen, J. M., Cutler, A. Butterfield, S. Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and cognitive processes*, 16(5), 637-660.
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor, Michigan: University of Michigan Press.
- Port, R., Cummins, F., & Gasser, M. (1996). A dynamic approach to rhythm in language: Toward a temporal phonology. In *Proceedings of the Chicago Linguistics Society* (pp. 375-397).
- Port, R. F., Dalby, J., & O'Dell, M. (1987). Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America*, 81(5), 1574-1585.
- 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.
- Takano, Y. (1995). *Gengo to Shiko*. In Y. Ohtsu (Ed.), *Ninchi Shinrigaku 3 Gengo* (pp. 245-259). Tokyo: Tokyo Daigaku Shppan-kai.
- Turk, A., E. & Shattuck-Hufnagel, S. (2000). Word-boundary-related duration patterns in English. *Journal of Phonetics*, 28, 397-440.
- Yoneyama, K. (2002). Lexicon and its stored representations for lexical access. *Journal of Phonetic Society of Japan*, 6(2), 23-34.