

日本語母語話者におけるP-centerの位置

石田, 英明
九州大学大学院比較社会文化学府

<https://doi.org/10.15017/26009>

出版情報 : 比較社会文化研究. 33, pp.1-11, 2013-02-15. 九州大学大学院比較社会文化研究科
バージョン :
権利関係 :

日本語母語話者における P-center の位置

The P-center locations of Japanese L1 speakers

イシダ ヒデアキ
石田 英明

1. はじめに

日本語学習者の発話においては、促音・撥音・長音といった特殊拍の習得が困難であることが指摘されている。これは学習者の多くが音節リズムあるいは強勢リズム言語を母語としており、目標言語（日本語：モーラ言語）の有標性が高いことが原因だと考えられている。日本語においては特殊拍が語の弁別にまで関わることから、この特殊拍の知覚と産出を含む¹「モーラ・リズムの習得」は学習者にとって非常に重要である。

これまでの日本語学習者や母語話者に対する日本語リズムにおける実験や調査では、主に音節の音響的 onset を基準にした子音長、母音長、音節長など持続時間を計測し評価してきた（戸田, 2003 など）。しかし、音響実験では音節 onset 間の時間長における拍の等時性は否定され（Beckman, 1982 ほか）、「その言語の母語話者が当該言語を聞いた時に感じる心理学的な等時性だ」（城生, 2011）と言われている。発話における拍を時間長で比較することの有効性に疑問が投げかけられている。とはいえ、多くの日本語教育の現場では「特殊拍を含む 1 音節は普通拍の 2 倍の長さ」だとして手拍子などを用いて指導している。学習者に対するリズム評価には母語話者教師による主観評価が用いられることも多く、数値的な客観評価は少ない。音節 onset を基準にした時間長における等時性が否定されていることが、学習者のリズム能力の数値化を困難にしている状況とも考えられる。数値的基準としては、子音と母音の持続時間の割合を測る RM (Ratio Measures)、言語リズム研究で利用される IM (Interval Measures) や PVI (Pairwise Variability Indices) を応用したリズム評価研究もあり、妥当性も示されている（木下, 2010）。しかし、人は子音・母音をひとまとまりにした音節を単位にリズム感知しており、この手法はリズムの生成・知覚を直接取り扱っているとは言い難い。子音長と母音長の比率を統計的に扱う手法のため、学習者に直接的、効果的なフィードバックを与えようかどうか疑問が残る。

一方、神経生理学的には、リズムとは連続的な音響的

な流れを心内ペースメーカーによって離散的な知覚に変換する働きと考えられる（小野田, 2004）。これは「感覚的处理モデル」が働く比較的短い時間に対する処理について述べられたものだが、この考え方を援用すると特殊拍の知覚は時間長の計測からリズム感知ポイントの「有無」の問題に変換されうる。さらに音響心理学分野では、音楽や言語のリズム認知はイベント（音事象）の中の Perceptual center (P-center) というポイントが関係しているという一連の研究がある（Morton et al., 1976 ほか）。それによると、人は P-center によって音の連続が時間的に規則的かどうかを判定しており、知覚と産出の両方に関わるといふ。この P-center を利用した言語類型研究としては、位相の概念を取り入れた Speech Cycling Approach (Cummins, 2002 ほか) があり、有効性が確認されつつある。つまり、日本人が日本語において感じる拍の等時性もこの P-center を基準に認識していることになる。従って、それぞれの音節の P-center が客観的に特定されれば、学習者リズムの数値評価はより直接的に顕示化することができるであろう。しかし現時点では、P-center はまだ解明されていない部分も多く、P-center 位置の推定モデルは統一されていない（Villing et al., 2011）。また、音一般の特性の中で語られるため、言語の違いによる変異は検証されていない。

そこで本研究では、調音運動や音響的变化をもとに日本語母語話者を対象とした普通拍音節の P-center 位置を調査し、日本語における P-center 位置推定モデルを提示することを目的とする。ただし、ここで考察する P-center は音全般に汎用しうるモデルではなく言語音声に限ったものとして議論する。さらに、日本語学習者の音声評価に応用すること自体が目的であるため、視覚的に確認可能な音響的特徴のどこに P-center の近似値が存在するか、という暫定的な位置の確認を目標とする。

2. 先行研究

2.1 P-center とは

P-center とは、音楽の音や発話音声において、「短い

音象 (event) の生起が知覚される仮定的な特定の瞬間」と定義され (Morton et al., 1976) リズムの知覚と産出に関わる。言語において話し手、聞き手に意識されているそれ自身に切れ目のない音(の連続)が音節であり(斎藤, 2007)、したがって1音節の中に1カ所P-centerが存在することになる。2つの音が「同期した」と感じるときは、その2つの音の音節 onset ではなくP-centerが同期しているのであり、また音の連続が規則的か規則的でないかを判定するのもこのP-centerによる。また、P-centerには「前後関係独立仮説 context independence hypothesis」が示されている。これは、P-centerの時間上の位置は音自体に固定されており前後環境に影響されない、というもので、その後の研究によって検証されている (Villing et al., 2011)。

2. 2 先行研究におけるP-centerの位置と、音響的特徴との関係

先行研究によって提案されてきたP-centerの位置は、主に「母音 onset」に関わるポイントと「強さ」に関わるポイントに大別できる。以下、主なものを概観する。

Morton et al. (1976) では音響的なマーカーの位置が候補としてすべて否定されたが、Fowler (1979) において聞き手が調音活動上のタイミングに関わる音響上の情報に基づいて等時性を判断していることが示唆され、Marcus (1981) において、P-centerは母音 onset を基準にするが子音長や母音長からの影響を受け、母音 onset よりも先行するポイントとなると結論付けた。しかし Scot (1998) によると、母音の調査では音節長の違いはP-centerにそれほど大きな影響を与えないことが報告されている。

一方、Gordon (1987) は Perceptual Attack Time という学術用語を使用し、合成された楽器音で音の知覚ポイントを探った。この研究によると聴覚神経細胞における知覚の閾の近似値にポイントがあり、特に音の「立ち上がり」の音の強さの傾斜と時間長が関係するという。音の立ち上がりの急峻なものは、その強さの Cue で P-center が決定されるが、立ち上がりが遅いものは他の特別な Cue (音響スペクトルや局所的なピークなど) に影響され易い、という結果が得られている。

ただし、Villing et al. (2011) によると、Marcus (1981) をはじめ現在提案されている P-center 決定モデルはどれも一致した位置を予測しない。

2. 3 調音運動と音響的特徴

Fowler (1979) において調音活動とP-centerの関係が示唆されていることから、ここでは語頭環境における日

本語普通拍音節の調音運動と音響マーカーについて整理する。

音節は母音を中心に構成される。日本語音節の開始部分のパターンはV、CV、CCV (拗音を含む音節) がある。一方、それぞれの音節末においては、母音で終了し後続音節に移行する軽音節のほかに、母音が長音化した「長音」、鼻音 [m, n, ŋ] が付加される「撥音」、後続音節の頭子音と同じ子音が付加される「促音」といった重音節、および長音+促音または鼻音、鼻音+促音、長音+鼻音+促音などの超重音節のパターンが存在する²。日本語母語話者においては核となる母音終了後に付加される特殊拍を後続リズムと認識することから、特殊拍以前の時間の中にP-centerが認知されることは明らかである。そこで、主要な日本語音節の語頭位置における開始部分の調音運動プロセスと、その運動に関連した音響マーカーについて、主に『音声の音響分析』(Kent, R. D., and Read, C. 荒井ほか監訳, 2000) および『日本語音声学入門 改訂版』(斎藤, 2007) を参考に概観する。

1) 単独母音音節

母音の開始は音声波形の振幅包絡の上昇に現れる。口腔内の形状変化はなくフォルマント・パターンは一定している。フォルマント・パターンは舌の位置を示し、すなわちF1は舌の高低を(高舌母音ほどF1が低い)、F2は前後を(後舌母音ほどF2が低い)示す。開始に声門閉鎖音がある場合、急激な振幅包絡の増加が破裂に似たスペクトログラムを示すが、フォルマント遷移は示さない。

2) CV音節 (CCV拗音音節を含む)

無声閉鎖子音 [p, t, k] では時間軸に沿って「閉鎖解放に伴う破裂(短い垂直パルスもしくは雑音パルス)」場合によって帯気音(雑音)が現れた後、「母音の調音位置への口腔の変化(F1、F2におけるフォルマント遷移)」を経てフォルマントは「母音定常部」に達する。

有声閉鎖子音 [b, d, g] においては「声帯振動の開始」「調音点の閉鎖(ギャップ)」「解放に伴う破裂(雑音パルス)」「母音位置への舌の移動(フォルマント遷移)」「母音定常部」となる。語頭では調音点の閉鎖は声帯振動の前に生じている。

無声摩擦音 [s, ç, h, ç, ç, ç] および有声摩擦音 [z, z] では「口腔内の狭めと呼気の乱気流(持続的雑音)」「母音位置への舌の移動(フォルマント遷移)」を経て「母音定常部」へと移行する。有声摩擦音では開始時の持続的雑音に声帯振動が伴うため、両者の違いは雑音部分の振幅の差である。

無声破擦音 [tç, ts] は「閉鎖の解放に伴う破裂(パルス

雑音)」「破裂に続く摩擦(継続的雑音)」「母音への舌の移動(フォルマント遷移)」「母音定常部」という流れで、有声破擦音 [dz, dz] は有声閉鎖音と同様、声帯振動の開始が先行し摩擦中も継続する。

接近音(わたり音) [j] は母音 [i] に、[w] は [u] に似た声道の狭めであり、開始のプロセスは母音の開始に準じる。まず [i, w] のフォルマント・パターンにおける「振幅の上昇」とともに「主母音へ舌の移動に伴うフォルマント遷移」そして「母音定常部」へと続くが、他の子音に比べ遷移時間が長いのが特徴である。

拗音は、先行子音の硬口蓋化によって母音遷移が [i] の形状から始まり主母音へと向かうため、遷移が比較的長時間になると考えられる。

鼻子音 [m, n, ŋ] は口腔での閉鎖があり、音節開始は

「声帯振動による鼻音マーマー(低周波の鼻音フォルマント)」「口腔における解放(母音フォルマント)」「舌の母音位置への移動(フォルマント遷移)」から「母音定常部」へと続くが、口腔の解放時に雑音は生じない。

弾き音 [ɾ] は「声帯振動の先行」「調音点への舌の接触に伴うギャップ」「解放に伴う振幅の増加」「フォルマント遷移」「母音定常」となるが、他の閉鎖音に比べ「閉鎖」時間が短いのが特徴である。ただし語頭の場合は比較的長くなると推測される。

これらの音節開始における順を追った調音プロセスと音響的变化のポイントは、表 1 のようにまとめられる。「○」はその運動が行われること「-」は行われないことを示し、従って○の位置が今回の実験における P-center 候補となりうる。

表 1 : 調音プロセスと音響的特徴

順	調音プロセス	音響的变化点	V	鼻音 CV	無声 CV	有声 CV
1	声帯振動	F0 onset	○	○	-	○
2	破裂、摩擦、接近	雑音 onset	-	-	○	○
3	母音	遷移 onset	-	○	○	○
4	母音定常部	F1F2F3 安定	○	○	○	○

注 1) 接近音においては、雑音は生じない。

注 2) 拗音において、母音 onset は [i] に近い母音のフォルマント遷移が始まったポイントと考え、[i] から主母音への遷移は母音中の変化と考える。

2. 4 先行研究の調査方法

先行研究の調査方法は主に、被験者が刺激音のリズムを調整する「調整法」と、刺激音に合わせてタッピングする「タッピング法」に分けられる。以下、これまで使われた測定実験の方法を (Villing et al., 2011) をもとに概観する。

2. 4. 1 音声ループによるリズム調整法

音声知覚において知覚閾を調査するため実験手法の「調整法」を利用した方法である。録音された 2 種類の音(ベース音とテスト音)の繰返し(ループ)が使われる。ベース音は時間的に規則的に配置されており、被験者はベース音の間に 1 つずつあるテスト音を等時間隔になるよう時間的位置を調整していく。最終的にベース音の P-center 間の中間点がテスト音の P-center 位置と判定される。ベース音としてテスト音自体を利用する方法のほか、一般的な音(参照音)を使う方法がある (Cooper, 1986 ほか)。これらは実験方法が理解しやすく簡単な装置でできることが利点であるが、言語音のベース音を採用する場合、基準となる P-center 自体が確定されていない

い問題点に加え、「調整法」は一般に微妙な主観的判断であることによる判定の困難さと妥当性の低さが指摘されている (『音声知覚の基礎』ライアルズ, 2009)。また、実験自体にかかる時間も長く被験者の負担が大きいと考えられる。

またこれを応用した方法もいくつか実施されている。まず、ベース音と同時に重ねるようテスト音を調整する方法 (Gordon, 1987) が使用されているが、聴覚マスキング効果や刺激融合効果などの影響で判定が困難であることが指摘されている。次に、「ベース音、ベース音、ベース音、テスト音」の連続を聞き、テスト音が遅すぎなかったか問う方法 (Fox and Lehiste, 1987) などもあるが、これは聴覚知覚において最後の音が過小評価される現象があることから妥当性が疑問視されている (Villing et al., 2011)。

録音されたテスト音を調整するのではなく、被験者が実際にベース音間にテスト音を産出する課題 (Fowler, 1979) や、産出とともに筋電図を用いる実験 (Fox & Lehiste, 1987) も試みられたが、産出に関わる調音運動の複雑性から音声の多様性が生じ、結果の正確性に保証

がないとの指摘がある (Villing et al., 2011)。

2. 4. 2 タッピング法

これは等時間隔で出されるテスト音の連続に合わせて被験者に指でタッピングをしてもらう方法である。タップされたポイントから反応時間等のバイアスを計算し P-center の位置が推定される。この方法は被験者の反応調査で意識的な判定を伴わないため、心理的負担は小さい。しかし、バイアス処理に関わる仮定 (バイアスは個人で一定である) が正しいという保証がない。この方法を改良した「位相修正反応法 (Villing et al., 2011)」も考案されている。ただし、タッピング実験には専門的な装置が必要となる。

以上、先行研究より P-center の位置及び調査方法を概観した。ただ計測から導き出された P-center 決定モデルはどれも違うポイントを導き出してしまい (Villing et al., 2007)、P-center は事象ごとに計測される必要があるという。しかし汎用されうる「学習者の直接的・客観的リズム評価」のためには、暫定的であれわかりやすい P-center 決定モデルが必要である。

3. 本研究における目的と仮説

目的)

本研究の研究課題は「日本語母語話者は、各種の日本語普通拍音節に対し、音節上のどの音響的マークに P-center の近似値を認めるのか」を明らかにすることである。

仮説)

カクテルパーティー効果等によって、言語音は一般音と別の処理プロセスの存在が指摘されており、言語音に特徴的な P-center が想定される。言語音における音事象の音響的過程は、声帯振動の生起、漸進的な振幅の増加 (強さの閾)、子音の口腔内運動による音響マーカー、母音生成のための舌の形状変化に伴う遷移開始マーカーがそれぞれ P-center の候補と考えられるが、知覚の局面での音響の振幅 (音量) や調音活動の面での舌の筋肉運動量を考慮すると「母音遷移開始マーカー (母音 onset)」が優位に選択され得ると考えられる。また、リズム認知と音節種認定を同基準で処理することができれば、情報処理資源の省力化が可能である。そこで、本研究では Marcus (1979) によって提示された母音 onset (フォルマント遷移の開始) に近いポイントに、日本語普通拍音節においても P-center が存在するという仮説を設定する。

4. 実験調査

特殊な実験装置を必要としない「リズム調整法」に準じ、かつ被験者の負担を軽減するため「恒常法³」を援用した実験方法で行う。すなわち、1種類の音節についてあらかじめ数種類の P-center 候補で調整した「ベース音-テスト音」の連続音ループを作成しておき、それらを選択肢として被験者が最も等間隔に配置されていると感じるものを1つだけ選ぶという強制的な選択方式で実施する。

4. 1 方法

4. 1. 1 被験者

九州/中国地方出身、29～48歳、男性3名女性7名の計10名。特別な音楽教育経験は2名のみが「なし」で残り8名は1年～13年の経験があった。クリック音における等時性判断では2名が正解を出さなかったが、データにおいて他者と大きな逸脱は認められなかったことから除外していない。実験調査実施は2012年10月1日に7名、11月15、16日に3名。

4. 1. 2 刺激音ループ

1) 参照音 (人工音)

本研究では、言語音によるベース音ではなく一般音の参照音を採用する。630Hz、315Hz、160Hzの3種の正弦波を単純に合成し、テスト音を録音した男性の声に近い音質・音高を作成した (音脈分凝効果⁴を抑えるため)。また参照音の P-center を推定ほぼ 0 msec とするため、envelope は立ち上がり最高で続いて急峻な減衰を持たせ、音長を 20 msec とした。振幅の最高点 (-1 db) は 1 msec 以内にあり、急激に減衰して 20 msec のポイントで消滅する。

2) テスト音

調査対象とする子音は調音法・調音点を網羅するものの数は最小限となるように選択した。子音に後続する母音は /-a/ (破擦音のみ /-u/) とした。具体的には、低舌広母音 [a]、有声硬口蓋接近音 [ja]、無声歯茎摩擦音 [sa]、無声声門摩擦音 [ha]、無声軟口蓋閉鎖音 [ka]、無声両唇閉鎖音 [pa]、有声両唇閉鎖音 [ba]、無声歯茎破擦音 [tsu]、歯茎鼻音 [na]、弾き音 [ra]、弾き音の拗音 [rja]⁵ の11種の音節とした。

テスト音の録音者は40代男性 (福岡出身)、使用した機材は富士通ノートパソコン FMV-BIBLO MG/B75、Sewell 製デスクトップ用マイク SPYKER によりデジタル録音 (周波数 44100 Hz)、編集は音響分析ソフト Praat (University

of Amsterdam製) およびフリーソフトSoundEngine Free (ver.4.41)を用い、分析にはSpeech Analyzer ver.3.0(SIL製)も利用した。

テスト音の規格は音節長を200msec(速度変化によって調整)、音高を約170~160Hzに調整し、音量も規格化した。しかしその他の音響特徴(intensityなど)は原音を維持している。

P-centerの候補ポイントは、基本的に表1(調音プロセスと音響的特徴)で示した○の項目とIntensityの頂点としたが、2つの候補の間隔が約10msec以下の場合な

どは前後いずれかの項目のみを候補とした。また比較的子音長が長いものは途中で候補を入れた。すなわち、「さ」⁶「は」「りゃ」は漸進的に増加する子音intensityが-30db程度に上がった所(スペクトログラム上の明確点)、「つ」は子音継続中一度段階的に上昇したポイント、「か」は帯気音のonsetを候補として設定した。最終的なP-center候補(選択肢)を表2に記す。数字は音節onsetからの時間(msec)を示している。音節欄は左から実際の実験順に並べられている。

表2 各音節のP-center候補

音節	さ	ば	あ	ぱ	りゃ	か	は	や	ら	な	つ
1.F0 onset		0	0		0			0	0	0	
2.子音 onset	0	↓		0	33	0	0	22	23		0
(破擦中摩擦 onset)											16
(子音中のマーカー)※	30				54	15	30				39
3.母音 onset	58	46	18	24	100	27	60	54	45	46	77
4.Intensity 頂点	80	88	58	↑	↑	68	95	83	78	75	97
(F1 安定)				75							
5.母音定常部	↑	↑		50	↑	↑	↑	↑	↑	100	↑

注1) 色付きのマスには該当候補がないことを示し、矢印は他の候補位置に接近していたため矢印先の候補にまとめたことを示している。

注2) 「あ」において、F1、F2がスペクトログラムで明瞭に分岐した時点を経済上「母音 onset」に当てはめている。

注3) 「ば」においては遷移が緩やかであったため、ほぼ遷移完了(5.母音定常部)と完全に安定したポイントを設定した。「りゃ」の母音定常部は100msecを超えていたので設定しなかった。

注4) 「や」における子音 onsetはF2F3がスペクトログラム上で明瞭に現れたポイントとした。

注5) ※子音中のマーカーは、スペクトログラムに何らかのマーカーと見られる変化が存在した場合に設定した。

3) ループの作成

河野守夫(1997)では、音声理解が可能な全体的知覚ができる話速を330msec以内の間隔としている。そこで、本研究では音節間隔を300msecで統一することとする。

まず、「参照音」―「テスト音」を両者の音節 onsetが300msec間隔になるよう配置し全体を600msecとなるよう無音部分を加えたユニットを作成、これを8回繰

り返したものを基本ループ(音節 onsetをP-center候補としたループ)とした。次に、各P-center候補が参照音 onsetからちょうど300msecの位置となるように位置をずらし全体600msecにしたユニットをそれぞれ作成、それらを8回繰り返した各候補のループを作成した。すべての候補ループの開始部分に参照音による速度提示リズム(600msec間隔で4回繰り返したもの)をつなぎ、選択肢を完成した。

図1 ループ音のモデル



注: 音節「あ」を例としてループの構造をモデル図化すると上記のようになる。図中の1マスは600msec、●は参照音で、平仮名はテスト音。●は0msecに開始し、テスト音「あ」のP-center候補ポイントが●のちょうど300msec後に位置するよう、一定に「あ」の音節 onsetをずらして選択肢ループを作成する。

4) 音節、ループの順序および配列

提示する音節の順序は、調音法の似ている課題が続かないよう配慮した。具体的には「か-な-さ-ば-あ-ば-りゃ-は-や-ら-つ」の順。各音内での選択肢はランダムに配置した⁷。

4. 1. 3 装置

Microsoft Power Point ver. 2007によって調査内容をデータ化し、被験者自身が操作してページを移動できるようにした。1ページに1音節分の選択肢ループをすべて並べ、選択肢の順序はランダムに配した。選択肢番号をクリックすれば音を聴取できるようにした。ヘッドフォンを使い、快適で十分な音量であることを被験者に確認した。回答は紙媒体の用紙を準備し、記入してもらった。

4. 1. 4 手順

被験者はパワーポイントの最初の画面が立ち上がっているPCの前に座り、回答用紙を受け取った。PCの操作方法の説明を受けたあと、各自のペースでパワーポイン

トのスライドを進めた。等時間隔については「左右の足でスムーズに歩いているように聞こえる音の連続。引きずったような感覚がしないもの。」と説明され、例示されたクリック・パタンの視聴によって理解を確認した。課題はまずクリック音による等時性判断課題、続いて音声によるP-center選択課題を行った。選択肢を選ぶ過程では、音の再生は判断がつくまで何度でも聞き返してよいこと、順番通り聞く必要はないことを伝えた。ただし、迷った場合も必ず1つ選ぶよう指示された。どの被験者も20～30分ですべての課題を終えた。

4. 2 結果と分析

4. 2. 1 実験調査の集計結果

実験調査の結果は表3に示すとおりである。調音プロセスを考慮した音響的特徴を左の欄に、音節の種類を上記し、表中の数字はその候補を第一に選んだ者の人数である。(なお表中の「-」は候補として設定されていないもの、↑あるいは↓は接近しているため分離して提示できず合わせて候補としたものである)。

表3. P-center候補と被験者によって選択された数

	りゃ	ら	ば	な	さ	あ	や	ば	は	つ	か	合計	割合
1.F0 onset	①	①	①	②	-	④	③	-	-	-	-	11	10.0%
2.子音 onset	6	5	↓	-	①	-	1	④	④	③	③	27	24.5%
※(破擦音第2子音)	-	-	-	-	-	-	-	-	-	1	-	1	0.9%
※(子音中マーカー)	2	-	-	-	2	-	-	-	1	4	4	13	11.8%
3.母音 onset	1	2	9	6	5	5	4	3	4	1	3	43	39.1%
4.Intensity 頂点	↑	2	1	2	2	1	2	↑	1	1	0	12	10.9%
※※(F1安定)	-	-	-	-	-	-	-	2	-	-	-	2	1.8%
5.母音定常部	↑	↑	↑	0	-	-	↑	1	↑	-	↑	1	0.9%
	合計											110	100.0%

注) 数字は第一候補として選んだ人数。丸数字になっているものは音節onsetにあたるものを選んだ人数で、計26名(23.6%)であった。濃い色付のマスは競合のない第一候補、薄い色付きのマスは競合のある候補。表は競合のあり様から並び替えグループ化し表示したので、表2とは並びが異なる。

まず全体合計数で選択された数を比較すると、最もP-centerとして選ばれやすいのが母音onsetで43(39.1%)である。次いで子音onsetの27(24.5%)、その後Intensity頂点の12(10.9%)とF0 onsetの11(10.0%)が並ぶ。F0 onsetと子音onsetのうち「音節onset」に相当する選択肢を合計すると26(23.6%)であった。

また音節種類別にみると、競合のあり方で以下の3種に分類できる。

1) 競合なく(あるいは他の選択肢との差が大きく)1

つの候補に収斂されているもの

①「ら」「りゃ」・・・子音onsetに収斂(これらは音節onsetではない)。

②「ば」「な」「さ」・・・母音onsetに収斂

2) 2つの競合する候補(音節onset/母音onset)に収斂するもの

①「あ」・・・音節onsetとF1-F2明瞭点(表では母音onsetに分類している)での競合。

②「や」「ば」「は」・・・音節onset(子音onset)と母音

onsetとの競合。

- ③「つ」・・・音節onset(破裂子音onset)と雑音中Intensityの頂点との競合。

3) 母音onsetまでの候補がすべて競合するもの

「か」・・・音節onset(子音onset)、帯気音onset、母音onsetで競合

以上の結果を概観すると、P-centerは母音onset近辺で知覚されることが多く、これは仮説を支持し、英語でのMarcus(1981)の調査と同じ結果が得られた。言語による大きな違いはない。ただし本実験においては競合する位置として、ほとんどの課題で音節onset(音響的開始点)が選ばれている。この位置は、先行研究においてまず否定されている位置であり、実験方法による何らかのバイアスの存在が推測される。また、母音onsetの例外として、弾き音「ら」「りゃ」においては子音onset位置で知覚されることが圧倒的に多く、音節onsetとの競合さえない。その他の例外として、母音onsetが選択されなかった「つ」と、特定の収斂がなされなかった「か」の問題がある。

4. 2. 2 分析・考察

1) 音節onset(音節頭における音響的onset)の選択

音節onsetはMorton et al. (1976)において等時性基準としてまず否定され、そのことがP-center研究の出発点となったものである。数字のoneからnineの録音音声音を音響的onsetの等時性で並べても知覚的な等時性が得られなかったことからP-centerの概念が誕生した。また今回の実験における選択肢では、音節onsetが「単独で」多く選ばれることはなく、必ず競合相手が存在するのが特徴である。従って、この実験結果には実験方法によるバイアスがかかり、言語音のP-centerとは別に音の感度として音節onsetが浮上した可能性が考えられる。

音節onsetには、有声の声帯振動(鼻音のmurmurを含む)と無声の阻害音の2種類がある。

ここで注目されるのは、声帯振動開始(F0 onset)がP-centerとして多数選ばれているのは「あ」「や」の2種のみで、声帯振動の後に阻害音のある有声阻害音「ら」「りゃ」「ば」および声帯振動(murmur)の後に口腔の解放がある「な」では、ほとんど選択されていないことである。これらでは、F0 onset後に起こる最初の口腔内運動(阻害、解放)マーカーが選ばれている。つまり、口腔内運動で生じた音色の変化が、F0 onsetの感知をマスクしているような現象だと捉えることができる。一方、声帯振動が後発する無声阻害音では、歯茎摩

擦音「さ」以外のすべてで阻害音onset(=音節onset)がP-centerとして選ばれている。これらの事実を鑑みると、「音節onsetが選択された」と見られる現象は、阻害、接近や口腔の解放といった「口腔内運動の音響マーカー⁸⁾」がP-centerとして選択された現象だと言える。

一般音のP-centerを調査したGordon(1987)は知覚の閾値が関わっていること、また何らかの音色変化なども影響を与えることを示していた。従って言語音においても各種の「音響的マーカー」と「振幅感知の閾」の両者によって競合する可能性が考えられる。さらに、感覚器の感受性は被験者の状態によって変化することが知られており、閾によって感知される場合のP-center位置は注意の度合いで変化すると考えられる。つまり今回の実験のように、極端に短く単純な参照音との素早い繰り返しによって知覚が単純化され、さらに「開始ポイントの判断」という知覚の先鋭化を促されることによって、限りなく音節onsetに近い音響マーカー(口腔内運動onsetマーカー)でのP-center知覚を促したと考えられる。

以下では、上述の口腔内運動onsetマーカーと競合した「母音onset」を中心に分析と考察を進める。

2) V音節「あ」

「あ」に関しては音色変化や調音位置の変化などがなく、P-centerは「知覚の閾」に等しい可能性が高いと考えられる。音節頭以外の候補ポイントとして、便宜上母音onsetとしたスペクトログラムでのF1、F2明瞭表示点(約-30dB)と、Intensity頂点を挙げたが、Intensity頂点は1名にしか選択されておらずその他に候補も立てていないので、広く音節onsetからIntensity頂点までの間のどこかにP-centerが存在する可能性が示されているに過ぎないと言える。母音のP-centerの正確な計測には、極限法、恒常法などを用いた閾値計測が必要だと考えられる。

3) CV音節

CV音節の特徴としては、口腔内運動onsetマーカーと母音onsetとの競合であった。「や」「は」においては、両候補の間にも1つ選択肢があるにもかかわらず、ほとんど選択されていない。つまり選択は2者に分断されていることから、「競合関係」の裏付けと考えられる。これはCV音節全般の特徴と仮定して議論を進める。

- ①「ば」「さ」および「は」「か」「つ」・・・フォルマント遷移の重要性

母音onsetに強く収斂されて口腔内運動onsetとの競

合が見られない「ば」「さ」には、典型的なP-center知覚に関わる法則性が観察される可能性がある。しかし「ば」は破裂開始と母音遷移開始の間隔が5 msecほどで候補を別立てすることができなかつたため、第一候補が破裂onsetか母音onsetかを区別できない。従って、「さ」においてのみ確実に母音onsetへの強い収斂があると考えられる。

「さ」の調音動作は舌尖を歯茎に寄せるのみで接することもなく、国立国語研究報告100(1990)の資料(p.146)における図では他の子音や接近音と比較しても動きは小さい。開始時の感覚的インパクトは呼気に伴う気流の乱れのみである。発生した雑音の振幅の増大は非常に緩やかで、それに比べ母音開始の音響的インパクトは非常に大きい。母音生成がループによって繰り返しかつ知覚されれば、子音中での変化の閾の感知は困難となることが予想される。このような差が母音onsetへの収斂を生んだのではないかと考えられる。

ところで同じ摩擦音の「は」には競合があり、音節onsetでのP-center感知も同程度になされている。「は」の雑音の周波数帯域は1000~4000Hz程度と「さ」に比べて低く、「さ」の摩擦音の方がよく感知される⁹はずである。ここで指摘できるのは、「は」は声門摩擦音であり口腔形は母音と同じで、調音運動全般に変化がないことである。「さ」に比べ「は」の雑音時スペクトログラムは明らかに母音「あ」に近い。フォルマント形状の「無変化、連続性」が母音onsetのインパクトを相対的に引き下げていると考えられる。従って母音特性のあるフォルマント形状は声を伴う母音onsetを指定するだけでなく、雑音を含めた広い範囲でP-centerの知覚に影響を与えている可能性が示唆される。

一方、無声軟口蓋破裂音の「か」については①音節onset、②母音onsetに12msec先行する帯気音onset、③母音onsetの3候補で競合が見られる。特に子音中の帯気音開始ポイントで選択者が微妙に多くなっているが、帯気音は口腔摩擦音として現れることから「は」との共通性が見られる。「か」において帯気開始ポイントで母音フォルマント形に先行するF2が現れたことで、実際の母音onsetとの連続性が発生している。帯気音は音量的には非常に小さいが、声の母音onsetを凌ぐP-center候補となった可能性もある。

さらに無声破裂音「つ」において、上述の流れから母音onsetが予測されるが、実際には母音に先行する摩擦音中のあるポイントがP-centerとして知覚された。録音されたテスト音において、摩擦音中のIntensityが1カ所で段階的に上昇しており、母音onsetの振幅包絡の上昇の方が明らかに顕著であるにもかかわらず、この摩擦

音中のポイントもP-centerとして知覚した被験者が最も多かった(母音onsetを選択したのは1名のみ)。これは「感知された変化のうち最も早いものをP-centerと知覚する」傾向がある可能性も指摘できるが、実験設定上の偏りが誘導した可能性がある。[t]の破裂の後の摩擦音と母音[su]は子音と母音で全く口腔内の形状が変化しない点において、「は」の発音プロセスと同じである。「つ」のスペクトログラムを精査すると摩擦音の生起は音節onset後約20msecのポイントであり、残念ながらこのポイントには候補を立てていない。候補とした39msecはIntensityが一段階上がって安定したポイントである。もし摩擦音生起時の20msec時も候補に含めていたならば、そのポイントがより多く選択された可能性も示唆される。

② 鼻子音「な」

上述の「ば」「さ」とともに母音onsetへの強い収斂が見られた。鼻音murmurの直後、口腔の解放が歯茎と舌においてなされ、この解放が母音onsetとなる。従って、スペクトログラムに見られる「フォルマント遷移開始ポイント」がそのまま「口腔内運動onset」であり、離れていれば競合する可能性がある両者が同一であることが、このポイントへのより強い収斂につながったと考えられる。

③ 接近音「や」・・・主母音の優位

子音[j]は母音[i]と口腔形状が同じで、主母音の[a]に素早く遷移する。この音の特殊性は子音が母音要素のフォルマント型を持っていることであり、語頭環境において口腔内運動onsetを目当てに知覚されるのなら[i]フォルマント中のIntensity知覚閾(便宜上、子音onsetとした)、主母音[a]フォルマント型を目当てにP-centerが知覚されるのなら母音onsetが選択されると考えられた。これは「音声上の最初に現れる母音的フォルマント型」対「音韻上の母音フォルマント型」という問題に変換しうる。結果は音節onsetと母音onsetのほぼ競合であった(子音onset<音節onset<母音onset)ため、通常P-centerは母音onsetで感知されると考えられる。従って、いわゆる主母音の生起がリズム知覚に関わっていることが示唆された。

④ 弾き音「ら」「りゃ」・・・口腔内運動onsetがより優位な音節

今回の実験で口腔内運動onsetが母音onsetよりも優位な結果となった音節に「ら」「りゃ」がある。弾き音を含む「流音」は一般に子音の共鳴性が高いことが特徴で、

母音の声道を形作りながら舌を打ち付けるので子音中もフォルマント型は明瞭に現れる。従って、母音の声道形によるフォルマントの出現によってP-centerを感知するという仮説が成り立つ。

また、調音運動のインパクトからも考察は可能である。この調音運動のインパクトが母音onsetを凌駕した可能性を示唆する。国立国語研究所報告100(1990)では、日本語[ra]の調音運動プロセスは舌尖をせり上げ、素早く歯茎に接触し、後続母音の舌の位置まで移動する。語頭では接触時間がより長くなると考えられる。破裂音の完全な閉鎖と単純な破裂とは違い、舌尖の筋肉による複雑なコントロールが必要となる。つまり子音調音動作における舌尖の動きのインパクトの方が母音の調音運動に伴う舌の移動よりも注意の向くポイントとなりやすいことが考えられる。ただし、スペクトログラムには調音運動の複雑性は顕示化され得ず、聴覚知覚における運動理論¹⁰(音声を知覚する際に調音運動のCueを利用しているとする)を前提としなければ説明はつかない。

5. まとめと今後の課題

今回、日本語母語話者は各種の日本語普通拍音節に対し、音節上のどの音響的マークにP-centerの近似値を認めるのか明らかにするため、日本語母語話者10名に対し実験調査を行った。その結果、以下の点が明らかになった。

- 1) 日本語普通拍音節で日本語母語話者を対象に調査したP-center(近似値)では、口腔内運動onsetとフォルマント遷移のonsetの2種の音響的マークが競合状態にあることがわかった。これは、言語音のP-centerと一般音のP-centerが異なる可能性を示している。すなわち、一般音のP-centerは感覚の閾値に関わり、言語音のP-centerはフォルマント遷移onsetを反映していると考えられる。
- 2) 通常的环境下では、当該音節の「主母音」におけるフォルマント遷移のonsetが優位に選択されると推定されるが、主母音の声道形を認知しうるフォルマントをも子音であれば、その口腔内運動onsetにおいてP-centerを感知しうると考えられる。また母音音節は最初から主母音が現れるため、そのP-centerは知覚の閾値に関わると考えられる。その値については今回の調査では解明し得なかった。
- 3) 「その音の生起した時間」を特定することに注意が先鋭化した実験的環境の場合、口腔内運動onsetの位置、すなわち何らかの音色変化が感知し得る時点が

P-centerとなる可能性がある。

今回の実験で得られたP-centerは語頭環境のものである。本来P-centerは環境に関係なく音自身に結び付けられた位置であるが、そのことを何らかの方法で確認する必要がある。特に母音が語中に来た場合、そのフォルマント遷移が音節マーカーとして機能しているのか、もし違う結果があるのなら、どのようにP-center位置の確定方法を修正すべきなのか、明らかにしたい。そのうえで、より直接的な日本語学習者のリズム習得の計測・評価方法を開発していきたい。

【参考文献】

- 小熊利江(2008)『発話リズムと日本語教育』風間書房
 小野田慶一(2004)「時間知覚の神経生理学的基盤」『Behavioral Science Research』43(2), pp.79-88
 柏野牧夫(2000)「音声知覚の運動理論をめぐって」『日本音響学会誌』62(5), pp.391-306
 河野守夫(1997)「リズムの知覚と心理」杉藤美代子監修『日本語音声2 アクセント・イントネーション・リズムとポーズ』三省堂, pp.91-139
 木下直子(2010)「韓国人日本語学習者の日本語リズム習得研究」早稲田大学博士学位申請論文
 ケント, R. D. and リード, C. 著, 荒井隆行・菅原勉監訳(2000)『音声の音響分析(第3版)』海文堂, Kent, R. D. and Read, C., (1992) The Acoustic Analysis of Speech, Singular Publishing Group
 国立国語研究所(1990)国立国語研究報告100『日本語の母音、子音、音節 - 調音運動の実験音声学的研究 -』秀英出版
 斎藤純男(2007)『日本語音声学入門 改訂版(第2刷)』三省堂
 戸田貴子(2003)「外国人学習者の日本語特殊拍の習得」『音声研究』7(2), pp.70-83
 ライアルズ, J 著, 今富摂子・荒井隆行・菅原勉監訳(2009)『音声知覚の基礎(第2版第2刷)』海文堂, Ryalls, J. (1996) A Basic Introduction to Speech Perception, Singular Publishing Group
 Beckman, M. (1982) Segment duration and the 'mora' in Japanese, *Phonetica: International Journal of Speech Science* 39(2-3), pp.113-135
 Cooper, A. M., Whalen, D. H., and Fowler, C. A. (1986) P-centers are unaffected by phonetic categorization, *Perception & Psychophysics* 39(3), pp.187-196

- Cummins, F (2002) Speech Rhythm and Rhythmic Taxonomy, *Speech Prosody*, pp.121-124
- Fowler, C. A. (1979) "Perceptual centers" in speech production and perception, *Perception & Psychophysics*, 25 (5), pp.375-388
- Fox, R. A., and Lehiste, I. (1987) The effect of vowel quality variations on stress-beat location, *Journal of Phonetics* 15, pp.1-13
- Gordon, J. W. (1987) The perceptual attack time of musical tones, *Journal of the Acoustical Society of America* 82(1), pp.88-105
- Marcus, S. M. (1981) Acoustic determinants of perceptual-center location, *Perception & Psychophysics* 30(3), pp.247-256
- Morton, J., Marcus, S., and Frankish, C. (1979) Perceptual Centers (P-centers), *Psychological Review*, pp.405-408
- Villing, R. C., Repp, B. H., Ward, T. E., and Timoney, J. M. (2011) Measuring perceptual centers using the phase correction response, *Percept Psychophys* 73 (5), 1614-1629
- Villing, R., Ward, T., and Timoney, J. (2007) A review of P-Centre models, In: Rhythm Perception and Production Workshop, Kippure Estate, Co. Wicklow, Ireland. (インターネット閲覧2012年12月1日 http://eprints.nuim.ie/1432/1/TWrrppw2007_models_2007b.pdf)

註

¹ 小熊利江 (2008) では、普通拍が特殊拍のように発話される「拍の増加」の誤りが多いことが指摘されており、単に特殊拍のみの問題ではない。

² 日本語母語話者は重音節に関わる長音、撥音、促音をそれぞれ1拍と認識するため、1音節で最高4拍までのバリエーションが存在する。例えば「モンスーンって」における「スーンっ」は1音節で4拍となる。

³ 「用意した数段階の刺激をランダムに繰り返し提示し、それぞれの刺激に対する被験者の判断を求める方法。」(ライアルズ, 2009)

⁴ 周波数の異なる2種類の音AとBが交互に比較的早い速度で連続する場合、A-B-A-B-A-B…という並びには聞こえず、AのグループとBのグループが分かれて聞こえるため、等時性が判断できなくなる現象。

⁵ 実際には [na], [sa] において、子音長を100msecほど継続させた音声も作成しテストを行ったが、今回の報

告目的にはそぐわないため取り上げていない。

⁶ 本研究では、一般的な音節については音声記号 ([sa] など) を用いたが、具体的に録音された音について記す場合は「さ」のように平仮名を用いることとする。

⁷ 2011年9月に7名の実験実施したのち、順序を一度改定した。11月実施の3名は最初の2課題「か」「な」と全く同じ課題を後半に再び組み込み、後半のデータの方を採用した。複数の被験者から「中途から感覚が先鋭化されたことが自覚され、前半の判断に自信がない」との意見が出たため、最初の2課題を練習セッションと位置付けた。また「な」においては刺激音の分析が不十分で、最初の7名の課題には母音定常部の選択肢を入れていなかった。後の3名にはこの選択肢を加えたが、それをP-centerとして選んだ者はいなかったため重大な齟齬は生まないと考えた。

⁸ これは頭子音に関わる解放、接近、狭め、接触などすべての口腔内で生じる調音動作を指す言葉とする。

⁹ 国際標準規格化された等ラウドネス曲線によって確認。

¹⁰ 「音声知覚の運動理論 (the motor theory of speech perception)」A. M. Libermanによって示された説で、音声の知覚は生成活動が大きな役割を果たすというもの。懐疑的な批評も多いが、近年神経生理学分野の進展で再び見直されつつある。(牧野, 2000) より

The P-center locations of Japanese L1 speakers

Hideaki Ishida

The moment of perceptive occurrence of a sound event is defined as “perceptual center” (P-center). So far, P-center locations have been measured experimentally in English utterances and in sounds of instrumentals, and researchers have tried to make some P-center models. However, the valid predictive model has not been established, and P-centers have not been verified for the Japanese L1 speakers. On the other hand, the identification of the P-center locations in the syllables is indispensable to grasp language rhythms directly. In learning the Japanese language, the acquisition of special moras is important, and the P-center locations seem effective indicators of direct and objective rhythm evaluation.

In this study, approximate P-center locations of ordinary syllables in the Japanese language were identified by using looped repetitions of reference-test sounds and forced choices on the part of ten Japanese L1 speakers. Acoustic markers visible with spectrogram were set as the choices.

Analyzed with acoustic information connected with articulatory timing, a competition of “an oral cavity event onset marker” and “a vowel onset marker” is found in a P-center of CV syllables. In precedent studies, the P-center of the syllable onset location was denied, and most of oral cavity event markers are located there in silent obstruent consonant. Therefore, it is thought that “vowel onset marker” is usually chosen predominantly, but “oral cavity event marker” can be chosen in the experimental environment. In addition, at the syllables with the “flip” consonant (which has a complicated articulation), the oral cavity event marker is chosen overwhelmingly. Therefore, it is suggested that the choices may vary for different phonemes. However, the P-center location of V syllables was not identified sufficiently in this study.