

二重母音と連母音の違いは何か？

—音節構造から比較する英語と日本語の二重母音—

大 高 博 美

—本稿の構成—

- I はじめに
- II これまでに提唱されている主な二重母音の定義
- III 音節の本質
 - 1. 音節に関する疑問
 - 2. リズム単位としての音節
 - 3. 音節長を計る認知上のメカニズム
 - 4. 母音と子音を音楽理論の観点から解釈する
 - 5. 音節の定義
 - 6. 音節とモーラ：音節量
 - 7. 音節の内部構造
 - 8. 分節
 - 9. 音節とプロミネンス
- IV 音節の新定義に基づく二重母音の考察
- V 英語の二重母音第二要素は単母音かわたり音か
- VI 日本語に二重母音はあるか
- VII 結語

I はじめに

「子音」「母音」と並んで、英語を学んだことのある者なら誰でも知っている音声学関連用語に「二重母音」がある。現在、アメリカ英語を例に挙げると、[aɪ], [eɪ], [ɔɪ], [aʊ], [oʊ], [ɪə], [ɛə], [ɑə], [ɔə], [ʊə] の10種が認められている（安井1992）。しかしこの用語は、実は、小野（2011）で指摘されたように、明確な定義の下で使われているわけではない。後述するように、これまで様々な音声学者がそれぞれに二重母音を定義しているが、どれも決定的な学説となるには至っていないのである。日本語の母音を取り上げてみても、「二重母音」と「連母音」（もしくは「母音連鎖」）の術語が常に明確に使い分けられているわけではない。例えば中條（1989, p.19）では、「[ai], [oi], [ui] のよ

うに二個の母音が密接に結びついて発音されるとき、これを二重母音または重母音という」と説明する一方で、別のところ (p.19 注2) では「[ai], [oi], [ui] などの二重母音は一つの母音として実現されることが少なく、二つの母音の間に音響上の境い目が認められる。このように、二つ以上の相異なつた母音が対等に連結したものを連母音といい、二重母音とは区別する」とある。つまりここで問題なのは、二重母音と連母音の定義がさほど明確でないという点である。連母音の定義に使われている「一つの母音として実現されることが少ない」、「二母音は対等に結びついている」、「二つの母音の間に音響上の境い目が認められる」などの説明は、正しさにおいて自明ではないからである。この説明を読んで何となく分かった気にはなっても、客観的な見地から読者を納得させる説明とはなっていない。実際、二重母音と連母音をソナグラム上でフォルマントの変化を基に比べてみると、残念ながら、「二母音間の音響上の境い目」の明瞭さにおいて両者に決定的な違いは見られないのである (加曾利 2013)。

二重母音とはどんなものであるかが依然としてはっきりしない理由として、小野 (2011) は次のように述べている。

そもそも二重母音って何だろうとか、二重母音の教え方が分からなくて困っているとか、そういう悩みを英語教師が発するのを聞いたことがない。学会発表や音声学・音韻論の論文であっても、二重母音そのものを正面から取り上げたものはほとんどない。これは恐らく音声学の専門家を含めた英語教師の間で二重母音について今更議論することはなく、自明の事柄として処理されているからだろうと思われる (p.189)。

「二重母音」は、通常、どの音声学者による場合も「音節構造」と「聞こえ」(母音四角形における舌の「高低」とほぼ同義) という二つの観点から定義される。例えば音節構造を基盤とする定義の例を挙げると、「二重母音は一つの音節を形成する」(Jones 1960)、「一つの音節内で途切れなく移行する一つの音」(Kenyon 1951, 竹林 1976)、「一音のように発音される二つの母音」(鳥居・兼子 1969)、「音節核に現れる二つの要素からなる母音」(栞矢 1976)、「母音連続の中で単一音節に収まるものを二重母音と呼ぶ」(窪蘭 2002)、「二重母音とは一つの母音と見なせるものである」(Ladefoged 2006) などである。よって、例えば日本語によく見られる母音の連続 (例: 家 [ie], 声 [koe], クエ [kue] など) は、例外的に一音節に収まるとされる [ai] や [au] を除けば、どれも二重母音ではなく連母音であるということになる訳である (窪蘭 2002, Vance 2008)。一方、上述のような音節に基づく観点だけからでは「二重母音」をうまく定義することができない。英語においても日本語においても、二重母音として認めがたい (もしくは存在しない) 母音連続があるからである (例: 英語 [ia], [ie], [ua], [oa]; 日本語 [ie], [uo], [ea] など)。つまり上の定義では、ある種の母音連続がなぜ一音節として認められないのかが分からないまま

なのである。その最大の理由は、「音節とは何か」が明確に定義されないまま二重母音が定義されているからである。

上述の瑕疵を補うために導入されたのが、もう一つの視点「聞こえ度」(sonority)である。つまり、二重母音の始発母音(第一要素)の聞こえは終着母音(第二要素)よりも大きいというものだが¹⁾、母音の聞こえは舌の高低に連動する(低舌化するほど聞こえは高くなる)ので、実際の定義においては始発母音から終着母音への「方向」が使われるときもある。例えば、英語の二重母音は「方向性」の観点から、「上向き二重母音」([ai], [ou], [ei], [au], [ɔi], etc.)と「中向き二重母音」([iə], [ɛə], [uə], etc.)に分類できるとされる(鳥居・兼子 1969, 柘矢 1976, 竹林 1976)。これにより、日・英両言語において、ある種の母音連続がなぜ二重母音として認められないかが明らかとなるのである。しかし、ここでも再び問題が生ずる。シュワーは中舌母音で [i] や [u] の高舌母音より聞こえ度は高いはずなのに、何故 [iə], [uə] などは二重母音として認められるのかという疑問である。

このように、音節内での聞こえ度の変化に着目する上述の定義も完全なものとはなっていない。では Ladefoged (2006) のように、聞こえの代わりに「プロミネンス (prominence: 卓立)」を使い、「二重母音の前半部分は後半部分よりもプロミネンスが高い」とすればここでの問題が解決するかといえ、そうでもない。置かれるプロミネンスの観点から、彼の定義により [ju:] なども二重母音と見なされる訳だが、これは本当に二重母音であろうかという疑問が拭えないからである。つまり、たとえ強勢が音節の出だしから始まるにしても、/ju:/ は頭子音としてのわたり音 [j] と核としての母音 [u:] から成る CV 音節とは考えられないのかという疑念である。

前置きが長くなったが、本稿の目的は、二重母音や三重母音などの重母音の本質を音節構造の観点から詳しく考察し、連母音との違いを探ることにある。

II これまでに提唱されている主な二重母音の定義

過去の二重母音研究の成果に関しては小野 (2011) が批判的に詳しい解説を加えながら紹介しているが、本稿でも主な研究者による二重母音の定義を簡単に概観してみたい。本稿のIV節で提案する二重母音の定義は斬新なアイデアに基づくものだが、どのくらい斬新かはこれまでの定義と比べなければ実感できないからである。

下に示すのは、これまでに提唱された多くの研究者による二重母音についての定義である。公表された年代順に、Kenyon (1951) から Ladefoged (2006) まで九例ほど挙げてみる。

1) 実は、この学説も完全に確立したものとなっているわけではない。Kenyon (1951) や Trask (1996) などでは、二重母音には二種あり、プロミネンスが V2 に置かれる /ju:/ なども二重母音と見なされる。

1. Kenyon (1951) : アメリカで出版された著名な音声学書の一つ
 - (1) 一音節内で途切れなく移行してゆく一つの音
 - (2) a. 下降 ai, au, iu, oi, ou, ei, ir, er, ar, or, ur
b. 上昇 wi, we, wo, ji, je
2. Jones (1960) : 権威ある英語音声学解説書の一つ
 - (3) 二重母音は一つの音節を形成し、その音節のプロミネンスは上がるか下がるかであって、上がって下がるというようなことはない。
 - (4) a. 下降二重母音 ei, ou, ai, au, oi, iə, εə, ɔə, uə
b. 上昇二重母音 iə, uə, ui, oi, oə, eə
c. その他の二重母音 oi, eə, oə
3. 鳥居・兼子 (1969) : 60年代に日本で発行された英語音声学書の一つ
 - (5) 二つの母音を一音のように発音し、第一音のほうが第二音より音量が大である。
 - (6) a. 上昇 ai, ou, ei, au, oi
b. 中向き e(ə r), u(ə r), i(ə r), ɔ(ə r)
4. 柘矢 (1976) : 鳥居・兼子 (1969) 以降の日本で出た本格的な英語音声学書の一つ
 - (7) 音節核に現れる二つの要素からなる母音
 - (8) a. 上向き二重母音 ei, ai, oi, au, ou (英米共通)
b. 中向き二重母音 iə, εə, uə (英のみ)
c. r 二重母音 ir, er, r, or, ur
5. Trask (1996) : 音声学・音韻論の辞書の一つ
 - (9) /ju/ や /ai/ のように、一つの核の中で、ある母音から次の母音に音質が滑らかに移行する音である。通常、どちらかの母音は他の母音よりもプロミネンスが大きい。/ju/ のように、プロミネンスの低い母音が第一要素であるとき (on-glide : クレシェンド二重母音) と、/ai/ のように逆の構造をもつ場合がある (off-glide : デイミニュエンド二重母音)。
6. 小泉 (1996) : 90年代に発行された言語学の大家による音声学書の一つ
 - (10) 母音が出始めとは異なる母音の位置で終わるものを二重母音と称する。
7. 竹林 (1996) : これも本格的な英語音声学書の一つ
 - (11) 二重母音は単に二個の母音が連続するのではなく、Aの母音の調音位置から出発してBの母音の調音位置へと切れ目なく滑らかに移行し、第一要素の母音と第二要素の母音との間には途切れがなく境界もあまりはっきりせず一つの音節を形成する。
8. 窪蘭 (1999), 窪蘭・本間 (2002) : どちらも詳しい音声学の書
 - (12) 二重母音と呼んでいるのは、(日本語の場合) 漢語や外来語に特に現れる [ai] や [au] などの母音連続で、一般言語学的には、開口度の大きな母音から小さな子

音（母音の間違いか？）へと転移する（1999, p.149）。

- (13) 関与する二つの母音の聞こえの差が大きければ大きいほどその母音連続は二重母音になりやすい。ゆえに、[ai] や [au] は最も二重母音らしい二重母音であり、逆に聞こえの差があまりない [eo, oe, iu, ui] などは二重母音になりにくい母音連続である（2002, p.13）。

9. Ladefoged (2006) : 世界的に権威ある音声学による音声学書の一つ

- (14) 二重母音とは一つの母音と見なせるものであって、その音質は変化するものである。通常、前半部分のほうが後半部分よりもプロミネンスが高い。
- (15) a. ai, aʊ, ei, oʊ, ju: (英・米)
b. iə, eə, aə (英)

上に挙げた様々な二重母音の定義を見てみると、みなどれも簡潔ではあるが、内容的には同じものが一つもないことに気づく。しかし、表現は様々であっても、二重母音の定義においてはほとんどが「一つの母音とみなせる」、「二母音間に切れ目がない」、「ある母音から次の母音に音質が滑らかに移行する」などの表現が使われている点では一致する。しかしこれらの説明は、二重母音を定義する上でまだ十分なものとは言えない。すでに言及したように、音節という概念を不明確にしたまま、これを二重母音の定義に援用しているからである。音節核の理解にしても研究者間で見方が異なるが、これもまた不完全なまま使われる音節の定義に起因するものである。例えば、/ju/ のような第一要素がわたり音でも二重母音として認める者もいれば (Kenyon 1951, Trask 1996, Ladefoged 2006)、認めない者もいる。また、/ej/, /ow/ のように、二重母音の第二要素はわたり音表記が相応しいと考える者もいる (Hudson 2000)。それもこれも、結局は、音節が生成・知覚される際に関与する三種の次元（時間、音色、強さ）を区別してその本質を理解しようとしなからである。音節は、次節で詳述するように、本質的には時間次元上で処理されるリズム単位である。音楽理論的に言えば、主音とその前後に付随する小音とから成る音群に相当する（詳しくは本稿Ⅲ.4節を参照）。しかしこれだけでは、聴者は発話者が意図した音節を知覚できない。耳に入った音群をリズムの把捉という意味付けのために分節したくても、依拠すべき音響上の手がかりが何もないからである。ゆえに、音節の生成にはどの言語においても、音素によって異なる、強さ次元における内在的差異「聞こえ度」が組み込まれる。音節ごとに聞こえ上でピーク（一つの山）が出来る²⁾ ことにより、聴者は音節を知覚する上で大変有効な音響の手がかりが得られるからである。音節が持っている聞こえ度配列原理 (Sonority Sequence Principle) が普遍性の高いものとして一般化されている

2) 例えば英語の 'trend' という語を発音すると、音節核となる母音 (/ɛ/) をピークとして聞こえの山が形成されるが、このとき発話者は母音部をその前後の子音部より意図的に強く発音しているわけでは決してない。その音節を構成する音素群は一リズム単位として一回の外向肺臓流気で [trend] という音声連続体に具現されるが、音声によって聞こえ度が異なるために、上述のような結果となるのである。

のは、この理由に因る。

Ⅲ 音節の本質

1. 音節に関する疑問

音節の定義に関しては、これまで、内容的に大きく異なるものが数多く提案されているというわけではない。音節の核となる母音 (V) とその周辺部 (従音) となる子音 (C) に着目して、「母音を中心とする秩序だった音のまとまり」(窪蘭・本間 2002) というのが最も一般的なものであろう。これに従えば、/ka/ (「蚊」) のような開音節は CV、/get/ (‘get’) のような閉音節は CVC と、構造を母音と子音でそれぞれ表すことができる。しかし、音節の核となるのは常に母音とは限らない。英語の二音節語 sudden (/sʌdn/) や little (/lɪtl/) の語末子音のように、子音は時に成節化することが知られている (成節子音)。しかしこのような子音は核音として扱えばよいので、語の分節にあたってさほど大きな問題とはならない。「音節は核を一つだけもつ」という音節の定義を破るものではないからである。一方、もう一つの事実「核部が二重母音や三重母音によって形成された複合体をなす場合 (小泉 1997)」では、どうだろうか。例えば二重母音を含む音節 /spaɪk/ (‘spike’) の場合、確かに小泉 (1997) の言うとおり、V1 は広めの母音 (頂) で V2 は狭めの母音 (従音) であるから、音節構造の原理 (第一要素は第二要素より聞こえが大きい) に沿っている。しかし、常にそうとは限らない。例えば、英音の二重母音 [əʊ] (coat) は V1 に通常では弱母音となるシュワー (自然母音、中性母音、曖昧母音) をもっている。このシュワーで始まる母音は、Gimson (1962) で初めて指摘されてからイギリス英語の二重母音として認められているのである (安井 1992)。

上述の、音節を一つの母音とそれに先行・後続する子音から成る音連続体と見る定義法は、形 (form) に着目して定義したものだが、基本的には聞こえの観点に基づいてなされたものである。つまり、音節ごとに母音を中心として聞こえの山 (ピーク) が一つだけできるというのが肝心な点である。結果、成節子音もうまく処理できるというわけである。しかしこの定義では、多音節語の分節がうまくいかない場合がある。例えば、この問題でよく引用される英語の二音節語 ‘extra’ (/ɛkstrə/) を取り上げてみよう。英語の音素配列論に従えば、/ɛk.strə/, /ɛks.trə/, /ɛkst.rə/ の三通りの分節法が可能である。さらには、摩擦音 /s/ が両音節に跨る両節子音と見なされた場合、/ɛks.strə/ の分節法もありえる。よって、実際には四通りあることになる。この事実が投げかける音節の本質への疑問は無視できない。音節は本当に我々が言語を生成したり知覚したりする際の基本単位なのだろうか。窪蘭・本間 (2002) では、音節が英語の強勢付与やリズム生成に深く関わることを指摘しているが、これは音節が言語の生成と知覚に深く関わる音声単位であることを主張するものである。音節とは、言語学的に、いったいどのような単位なのだろうか。次節で考えたいのはこの点である。

2. リズム単位としての音節

本稿では、リズムを時間次元上で繰り返される複数音の束（パターン）、もしくはこのパターンが繰り返される現象、と定義したい。ここで時間次元とは、長さの概念を帯びる二次元空間のことである。ユークリッド幾何学的に言えば、点の集合体である。西洋音楽におけるリズムでは、この時間次元上で任意に決められる形（例えば二音から成るパターンでは「長・短」「短・長」など）以外に音の強弱も必須要素として加わるが、明確な拍子性をもたない邦楽（そして日本語も）ではその限りではない。つまり、強さ次元はリズム生成に必須の要素ではなく、リズム構成に必要なものはあくまで長さの概念であることが分かる。だからこそ、我々はまれに音とは関係のない動きにも（例えば視覚的に）リズムが感じられるときがあるのである³⁾。

では、我々はどのようにして「線」の長さを測る（認知する）のであろうか。数学では、線分 AB が与えられたとき、この線の両端を点 A, B とみなし、その間の距離を測ればよい。原理的には、時間の長さ測定法も同様のはずである。言い方を換えると、生成されたある音声の長さを計るということは、この音声の出だしと最後に二点の存在を認め、その間の距離（時間）を計るということである。リズムの生成・知覚とはまさにこの知的営みのことをいうのである。よって音節とは、発話者側に立てば、一つの線として発した音（群）であり、聴者側に立てば、聞こえ度における変化やリズム単位を手がかりとして一まとまりの線として捉える物理的音（群）のことである。例えば、日本語話者が一音節語「酢」(/su/) を生成するとき、それはモーラというリズム単位に基づいて一音節の物理音（音声）に具現される。この時、初めの点は /s/ であり、最後の点は線概念を帯びている核（母音 /u/）の最後の点である。換言すると、音節を構成する子音は点概念を帯びており、核となる母音と成節子音は線概念を帯びていると言える。つまり音節とは、線概念をもつ核の前後に点概念をもつ子音が集まってできた音の集合体であると定義できる。点は幅のない概念なので、一音節中に無制限に集うことができる。しかしこれはあくまで理論的な（認識上の）話であって、自然言語では自ずと制限される（例えば英語の頭子音数は最大三個、尾子音は四個まで）。そうでなければ、聴者は音連続を一音節として把促することが難しくなるであろう。認識する上では点として存在する子音も、具現化の際、物理的には長さをもつからである。繰り返すが、子音は音節中で点として存在するので、摩擦音や接近音のような継続音だけでなく、閉鎖音や破擦音のような非継続音でも子音となる資格をもつ。また、/ts/, /tʃ/ のような破擦音が物理的には二種類の要素からなっているも、単位として一つの子音と見なされるのも上の理由による。一方、核音は初めから線概念を帯びた存在なので、継続音しかこの位置を占める資格がない。典型的には母音であるが、比較的聞こえ度が高く継続性のある子音 (/l/, /m/, /n/, /z/) も可能であ

3) 例えば、車を運転していて道路沿いにある物体（電柱や標識など）が定期的に視覚に入り後方へ流れていくように見えるときなどである。

る。ここで聞こえ度の高い子音が資格を得るのは、聴者の便（音節を知覚するための音響的手がかり）が考慮された結果である。

音節は一つの線（核）とその前後に集まる0個から数個の点（子音）から成ると言えるわけだが、その本質を理解するにはもう少し詳しい説明が要る。我々は、点と線を異なる概念として観念上では理解可能だが、物理的にはこれらはどちらも存在することができない。点は幅も長さも持たない理想的な粒子であり、線は長さのみで幅のない概念だからである。よって今、誰かが平面空間上に筆記用具を使って点を表したとすると、物理的には長さも幅もあるものとしてしか具現できないのである。幅のない点が集まって線ができると考えるユークリッド幾何学における定理は、深い内容をもった問題である。長さが0である点をどのように集め並べたら長さ1の線分が得られるかなどということは、我々の理解の他にあっていい（志賀1992）。ともあれ、子音が音節中で長さを持たないと仮定すると、なぜ頭子音の数が音節の重さに関与しないのかが説明可能となる。何個集まっても意識の上では（音韻論的には）長さを持たないからである。CVもしくはCVCから成る音節を長く伸ばして発音する（もしくはテンポの極めて遅い歌曲として歌う）と、前者は最終的にVのみとなり、後者においても長く伸ばされるのはVのみである⁴⁾。理由は、繰り返すが、核のみが初めから線概念を帯びた構成要素だからである。

次に、ではなぜ尾子音は頭子音と異なる振る舞いをする（つまり音節の重さに換算される）のかという難問に答えなければならない。筆者が現時点で考えている解答は、音節を生成（具現）する際のスピード鈍化に基づく必然的な結果に基づくものである。音節を生成するというのは、長さを意識的に計るという知的行為が基本となるので、決められた長さの線分を定規で引くことに似ている知的作業（調音作業）である。例えば、ペンで5センチの長さに線を引くとき、出だし（数直線上で0）からは速く手を動かしても、決められた長さを意識すると、後方へ向かうほど自然に手の動きは遅くなるであろう。そうでなくては決められた長さに正確に線を引くことはできないからである。音節を生成する際にも同様のことが言える。違いは、「線を引く」ための手段が手の動きによるものか調音器官の動きによるものかという点だけである。すべての言語は決まったりズムを持ち、音節の長さはこれに基づいて、一つ一つ計られて具現される。つまり、音素を音声に具現する生成活動において、音節の出だしの調音運動は通常のテンポでは速く始まるが、音節末に向かつてはゆっくりにならざるをえないと考えられるのである。音素から音声への具現スピードが遅くなるということは、認識上の小さな点が物理的により大きな幅をもつ音声として具現されるということである。ゆえに、理論的には、すべての子音は幅をもたない点として音節中に存在するにしても、必然的に頭子音と比べて「長い」物理音となってしまう尾子音は音節の重さに関わるのである⁵⁾。音節を理解するためには、このように、発話

4) ただし、ゆっくり発音すると調音器官の運動も遅くなるので、物理的には子音の長さも長くなる。

5) 例えば、ラテン語のアクセント規則において、語末から二番目の音節が重音節（CVC, CVV）のときは強勢が

者と聴者の両方面から考察することが重要である。尚、尾子音と音節量の関係は、後のⅢ.6節でさらに詳述する。

3. 音節長を計る認知上のメカニズム

数学における線分 AB の長さ測定の幾何学的メカニズムと言語学における音節長測定の認知メカニズムがもし同じ原理に基づいてなされるものだと仮定すると、音節はその際に出だし部（起点）と括り部（着点）に点が意識されるはずである。大高（1988）ではこれらの二点を「出だし心理アクセント」「括り心理アクセント」と呼んだが、本稿ではそれぞれ、単に「起点」「着点」と呼ぶことにする。要は、ある音節の長さが発話者もしくは聴者によって計られる際には時間次元上でその起点と着点が必ず意識（認知）される必要があるということである。図1は日本語と英語において、それぞれの単音節（日：/ka/, 英：/strikt/）の長さが意識される際に置かれる起点と着点の位置を示している。太線（—）は核部（母音）の線概念を、そして円（○）は子音の点概念を示している。記号 [▽], [▽] はそれぞれ音長測定のための起点と着点を示している。/ka/ のような開音節の場合、母音の最後尾に着点が置かれる。この部分は、子音同様に点として意識されており音節の最後尾に位置するが、尾子音と区別するために記号の○は使わないことにする。

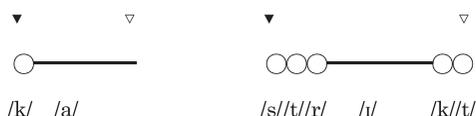


図1. 音節長測定のメカニズム：例 日本語 /ka/ (左) と英語 /strikt/ (右)

音節拍リズムをもつ言語⁶⁾においては、基本的に上述の計測メカニズムが採用されると考えられるが、モーラ拍の日本語やストレス拍の英語においては常に「付加リズム」が採られるわけではなく、もう一つ別の計測方法が採られる。「分割リズム」（もしくは「拍節的リズム」）のための音節長計測のメカニズムである。高名な音楽理論研究者であった Sachs (1953) は、音楽で採用されるリズムはすべて「付加リズム」（もしくは「自由リズム」）か「分割リズム（分節リズム）」に分類されると述べたが、私見では、言語のリズムにおいても同様である。

付加リズムでは様々なリズム単位が自由に連結されてパターンが出来上がる。例えば、/♪♪♪♪/ というパターンでは、八分音符で表された長さをもつ音が四回連結されている。また、/♪♪♪♪/ では音長比が1対2の関係にある音が交互に連結されている。つまり、付加リズムでは基本となる様々な音長単位が自由に組み合わせられて様々なリズム

ここに置かれるが、軽音節 (CV) のときは強勢位置が一つ前の音節に移動する。

6) 例えばフランス語やイタリア語などのロマンス系言語、中国語などもこの範疇下。

パターンが作られるのである。このことは、色の異なるテープが連続して繋がっている様子を思い浮かべると理解しやすい。この連続体中に切れ目（ポーズ）はないが、異なる色のテープごとにできる境界は明瞭であろう。それぞれのテープ（線分）の始まり（起点）と終わり（着点）がはっきりしているからである。日本語において二モーラフットを含まない語の場合（例：/hi.to/ 人）や、英語においてあたかも音節拍のように産出される特殊な文の場合（例：Tom told Sue three bad lies!）は、極めて付加リズムに近いものとして処理される可能性が高い⁷⁾。すべてのフットに一音節の内容語が当てられているからである。尚、リズムがテンポから独立するためにはある種のリズム単位が等時性を持つ必要があるのだが、音節拍ではこの役割を音節が担っている。

一方、分割リズムで等時性を担うのは「拍」（ビート）で、英語のリズムで言えばフットである。このリズムの生成においては、音節拍やモーラ拍におけるビートよりも比較的長い拍が与えられ、それが複数個の音（英語の場合、強勢音節と弱音節）で分割されるのである。図の2と3は分割リズムにおける音長計測メカニズムを図示したものだが、線分ABを二等分する際に使われる計測アクセントの数が四つではなく三つである点に注意されたい。二線を付加的につなごうとする場合とは異なり、一線を二分しようとするときには一点を間に置けば十分だからである（下図では記号 [◇] で表示：以下「中点」と呼ぶ）。換言すれば、その中点は先行音節の着点と後続音節の起点の両方を同時に兼ねているとも言える。図2での中点位置は、全体の二分の一のところに置かれている。

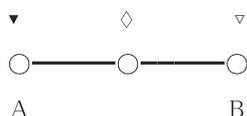


図2. 分割リズムにおける音長測定のメカニズム

さらに、例えば英語の六音節語 ‘rècommendation’ の場合は、生成もしくは知覚する際に使用されるリズムパターンは下に図示されるようなものになると考えられる。

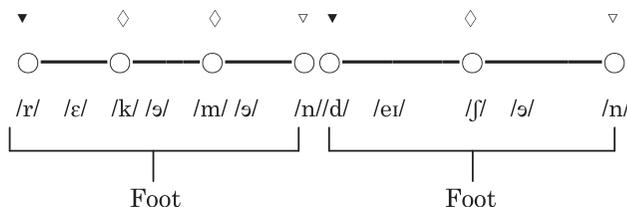


図3. 英語の語 Rècommendation がもつリズムパターンの構造

7) この文を実際に英語のネイティブスピーカーに発音してもらい録音し音声分析機に掛けてみると、確かに音節拍に近いものであることが分かるが、フランス語に見られるほどの等時性は見つけられない（大高 2012）。

図3が示すのは、この語は二つのフットを構造にもち（三音節から成る re.co.mmen と二音節から成る da.tion）、互いに付加リズムで繋がれているということである（つまり、起点と着点がそれぞれ二個ずつ使用されている）。結果、等時性はこれらのフット上で具現される。

4. 母音と子音を音楽理論の観点から解釈する

音楽における主音と小音（装飾音）の関係は、まさに言語の音節における母音と子音の関係に相当する（大高 1987）。唯一の違いは、前者が高さ次元と時間次元から成る空間上に具現される現象であるのに対し、後者は音色次元と時間次元上に具現される現象である⁸⁾という点だけである。図4は音楽における装飾音（上段）と実際の奏法（下段）を例示している。装飾音には単純装飾音と複合装飾音の二種があり、図では上下に分けて例示してある。長前打音は短前打音よりも心持ち長めになるように演奏せよという意味だが、共に独自の音価（長さ）は持たず次の主音に繰り入れられる⁹⁾。まさに音節中の子音と母音の関係と同じである。小音の数が一つの単純装飾音だけでなく複数の複合装飾音が存在する点も子音同様である。しかも、主音の前と後ろに付くことができる点も、子音における頭子音と尾子音の関係を想起させる。尚、音楽理論において主音（音符）は後続の異音程の主音とどのように連結されるのかという素朴な疑問があるが¹⁰⁾、これも本稿で提案した起点と着点に依る計測メカニズム原理を使えば理解できるようになる。



（画像の出典：<https://ja.wikipedia.org/wiki/%E8%A3%85%E9%A3%BE%E9%9F%B3>）

図4. 音楽における装飾音の例

- 8) 音素は音色次元上のみで規定された単位で、その音素体系にあっては長さ次元から独立している。そしてこれが時間軸上に具現されるとき音声（物理音）となるのである。
- 9) 短前打音が使われるようになるのは古典派以降（バロック時代とロマン派の間に位置し、百年記で言えば18世紀中庸から19世紀初頭に及ぶ時期）で、前打音が拍の前に出されることもあった（新音楽辞典 1977：音楽の友社 p.326）。この奏法の場合、勿論、装飾部の音符はあたかも弱起の曲における始まりのように後続の主音の音価から独立しているといえる。アクセントをどちらに置くかは楽器の種類（弦、鍵盤など）や時代そして演奏家の解釈にも依る。
- 10) 例えば、ド（↓）とレ（↓）の二音が連続して演奏されるときこれら二音は理論上どのように繋がるのか、つまり境界が必ずあるはずだが、その部分はどうのように接合されるのだろうかという疑問である。

ここで音符を使って、前述の‘extra’における曖昧な分節法の問題を再び取り上げてみたい。なぜ複数の分節方法が可能となるかが理論的によく理解できるようになるからである。‘Extra’ (/ɛk.strə/) の可能な分節法が一つに限らないのは (/ɛk.strə/, /ɛks.trə/, /ɛkst.rə/)、結局、どの分節法も英語の可能な音素配列論とリズム構造 (♩) に違反していないからである。つまり、二つの核 (/ɛ/ と /ə/) の間に挟まれている三つの子音 (装飾音: /k/, /s/, /t/) がどちらの母音 (主音) についても全体の二分割リズムに変化はないからなのである。繰り返すが、音節中の子音は音楽の装飾音同様に音価 (ここは長さのこと) をもっていないからである。

5. 音節の定義

音節に関するここまでの議論をまとめる意味で、本節では音節の定義を試みる。音節の本質は、すでに言及した通り、時間次元上でリズム単位として機能する点にある。音楽で言えば、主音を中心としてその前後に装飾音を従えることのできるリズム単位がこれに相当する。ただ、音楽の場合は時間次元と高さ次元での空間で規定されるが、言語の音節の場合は時間次元と音色 (音素) 次元での空間で規定される。言い換えると、音節とはその出だしと最後尾に音長計測を目的とする「点」が置かれた音もしくは音群のことである。例えば、英語の‘she’や‘strict’のような単音節語を単独で生成する際 (もしくは日本語の開音節から成る語を生成するときも)、ここでの音節は長さの計測に起点と着点が使われる。一方、英語の多音節語や日本語の閉音節から成る語を生成する際は、二音節を分ける中点は起点でもあり着点でもある。便宜上、前者のタイプの音節を自立音節と呼び、後者の従属音節と区別することにする。「従属」というのは、後者には分割リズムが関与するので物理的には決してポーズを使っての分割はできないという意味である。音節に挟まれた環境にある子音 (特に摩擦音) が両節性を示すときがあるのは (例: /sevən/ → /sev/ + /vən/)、このためである。

音節の最も基本的な本質は上述の通りだが、これはあくまで時間次元上だけでの定義であって、実際にはさらに聞こえ (強さ次元) も関与する。これがないと、聴者は発話者が生成した音群からリズムを聴き取る (つまり、聞こえの山を手がかりに長さのまとまりである音節に分ける) ことが不可能だからである。このことから、どの言語の音素配列論も発話者と聴者の両サイドの視点から双方向的に考案されていることが分かる。よって、言語ごとに音素配列論に違いが見られたとしても全体的には普遍性も認められる¹¹⁾。

6. 音節とモーラ：音節量

幾何学で使われる点と線概念に基づいて上で論じた音節の内容は、すべて理論上の話

11) 例えば、子音は音韻論的に長さをもたないので、/pm/などの同器官的子音連続や/sθl/-のような継続音の連続は避けられる可能性が高い。話者にも聴者にも点としての生成と認知が難しいからである。

しで、抽象性の高い説明となっている。幾何学での知見に基づいて点が集まって線が出来るといっても、その実際の長さ（物理量）は、長さの単位を決めないことには決まらないからである。例えば、10センチは1メートルより長いとは主張できないし、10m²と10mは比較にならない。比較する単位が同一ではないからである。音節の長さ比較においても同様のことが言え、私見では、この際に必要となる単位が「モーラ」である。モーラを使うことによって、発話テンポから自立した音節の長さ（内在時間）が比較できるようになるのである¹²⁾。ただ、モーラは長さを計る単位であるが、日本語と英語では使われ方がやや異なるので注意を要する。これは、リズム上で等時性の現れ方が日本語と英語では異なるからである。

日本語のモーラは、例えていうならば、距離を簡易に計るときに利用される足長（指先から踵までの長さ）のようなものである。実際、英語にはfootという距離の単位（12 inches）が存在する。人により実際の足長には個人差があるにせよ、通常、三足分の距離は二足分のものより長いと主張できる。よって、例えば「ロンドン」と「東京」は共に二音節語だが、四モーラから成るので「長崎」「広島」と同じ長さになるのである。これにより、日本語の音節には一モーラから成るもの（CV）と二モーラから成るもの（CVC, CVV）の二種があることが分かる。ただ、リズム単位としては、どちらも独立した長さ単位（二モーラフット）として機能するので、発話者もしくは聴者がCVN音節の長さを計る際には、認知上、先の図2の構造で示されるような計測方法が採られる¹³⁾。

ところで、伴奏なしの歌を極めて遅いテンポで歌うとき、我々はどうのように音節長を計っているのであろうか。例えば、テンポの遅い曲としてよく知られている日本の国歌「君が代」の歌唱を取り上げてみよう。この曲は四分の四拍子で、四分音符が通常一分間に50から70の間で演奏されるが、50というのはいかにゆっくりのテンポとなる。四分の四拍子であるから、勿論、歌う人は四分音符を一拍としてリズムを取るわけだが、このテンポの場合、一拍は1,200msとなる。この時、歌い手は無伴奏下でどのようにして拍の均一性を維持するのであろうか。強勢アクセントを持たない日本歌曲は無伴奏だと拍子が不明瞭になりやすいので、尚更である。

12) 先の脚注2でも指摘したが、音節の長さは発話のテンポにも左右される。例えば、英語においてゆっくり発音した'sit'は通常の'seat'よりも長くなるが、だからといってCVC音節はCVVCよりも長いとは主張できない。両音節をモーラ数で比べると、後者（超重音節）は前者（重音節）より一つ多く、実際にやや長くなる。ただしモーラ言語のように二倍長くなるわけではない。

13) 引音Rを含むCVV音節（例：香 /koo/）においては、意図的にゆっくり生成しない限り、図1に示される計測メカニズムが採られるものと思われる。一方、促音を含むCVQの場合は、構造が単音節（CV）+ Qであって、一つの音節とはなっていない。促音は、リズムの観点から言えば、音楽理論でいうところの休止（rest）としての機能をはたしているだけで、表層的に二重子音となるのは逆行同化によるものである（Otaka 2013）。もっとも、Qをゼロ音素とみなせばCVCと同様に扱える。

君が代

詩 古今和歌集 作曲 林 広守



実際に自分で色々なテンポで歌ってみれば気づくことだが、基本となる拍の長さ（ \downarrow ）は発話時のモーラ（ μ ）を等しく整数倍することで計っているのである¹⁴⁾。言語のモーラ長は、厳密には個人差があるが、通常は120ms（早口の人）から140ms（ゆっくり話す人）内に収まる。この長さは、調音器官を動かすときの筋肉運動のスピード感に結びついて、経験的に学習される。だから、発話のスピードに個人差はあっても、個人レベルでの発話スピードはほぼ一定する¹⁵⁾。

では、英語による無伴奏下での歌唱はどうか。英語の場合も、テンポを遅くすると調音器官の動きが比例して遅くなるので音節全体が間延びする。子音音素も母音音素も長く具現されるのである。しかし音節構成上は、あくまで点は点、線は線として機能している。つまり、基本的には母音部のみが音長調整の機能を担うのである。音楽で言えば、主音がこれに相当する。では、どうやって曲のリズムを維持しながら音節を等分に伸長できるのかといえば、拍子を構成する拍（ビート）の長さを伸ばすことによってである。西洋音楽では小節（measure）ごとにアクセント（強勢）が置かれるために拍子性は明確で、ビートが数えやすい。しかし、英語話者は、歌唱の場合とは違って、発話の場合には極端に遅いテンポで話すことは苦手である。歌唱におけるビート（拍）と発話におけるビート（フット）は似て非なるリズム単位だからである。前者は付加リズムであるのに対し、後者は分節リズムである。

ここからは、主題を歌唱における音節長測定法から言語における音節長測定法に戻そう。英語のモーラの役割についてである。英語のモーラも音節の長さを計る単位であるが、日本語のように決まった物理的長さ（等時性）をもっているわけではない。日本語のモーラが時に「音声的モーラ」と呼ばれるのに対し英語のモーラが「音韻的モーラ」と呼ばれるのは、このためである。英語のモーラは、実時間レベルで等時性をもつ単位ではないが、音節と並んで有用な単位である。アクセント規則をはじめとする音韻規則に深く関わっているからである。

英語における音節の物理的長さは、等時性を強く帯びるリズム単位であるフット（もしくはISI: Inter Stress Interval）がいくつの音節で分割されるかで相対的に決まる。例えば、先の図3で取り上げた英語の二フットから成る五音節語 ‘rècommèndation’ におい

14) 整数倍するといっても、実際に可能なのはせいぜい2~4倍であろう。よって、無伴奏下で極端に遅いテンポで歌うことは、通常、不可能なのである。

15) 発話のスピードは、勿論、様々な要因によって変化することが考えられる。例えばその要因には、驚き、悲しみ、疲労感などの感情が考えられる。

ては、最初のフット (rè.co.mmen) 中の音節は次のフット (dátion) 中の音節よりも相対的に短くなる。前者のフットは三音節から成るので全体が三連符のように発音され (一拍が三分割される)、後者のフットは二分割されるためである。

音節は音節量の観点から、次の三種類に分類される。尚、ここで頭子音数は、英語を例に、便宜上三個としてある。

- a. 軽音節 (一モーラ) : V, CV, CCV, CCCV
- b. 重音節 (二モーラ) : VC, CVC, CCVC, CCCVC, VV, CVV, CCVV, CCCVV
- c. 超重音節 (三モーラ) : VCC, CVCC, CCVCC, CCCVCC, VVC, CVVC, CCVVC, CCCVC

既に言及した理由により、頭子音は音節量に関与しないが、核中の母音数 (重母音) とその後の尾子音数は関与する。その理由は、すでに先のⅢ.2節で言及したが、理論的にはまだ不十分であるので、下でさらに解説を加える。

時間次元の上だけで音節を定義すると、「初めから線概念を帯びた存在の核 (単母音もしくは重母音の第一要素) の前後に点概念を帯びた存在の子音が数個 (もしくは0個) 付属してできる音群」となることは先にも述べた。点概念を帯びたものは幅をもたないので、実際にはそれが物理量をもつ形に具現されたとしても、観念上は (認識する上では) 長さが0なのである。頭子音同様に、尾子音も同じく点なので理論的には長さをもたない。しかし音節の生成に際し、音素群が音声群に具現されるスピードが後方に向かって遅くなるので、結果、尾子音は物理的に頭子音より長く具現されやすくなる。そして物理的な「長さ」が意識されるようになると、発話テンポが同じであれば、CVCはCVよりも「重い」と認識されることになる。ここで「重い」というのは、認識上の単位としては違いがなくとも (ここでは共に一つの線) 物理的長さでは勝っているという意味である。更に敷衍すると、点を紙の上に極細のペンで具現しようが、太いマジックペンで具現しようが、観念上は両方「点」と見なしうるが、物理的には後者の「点」の具現形の方が前者と比べて大きい (「重い」ということである。音声の生成と認識においては、観念的なレベルと物理的なレベルの両方が双方向的に関与する¹⁶⁾ ので、外界からの情報 (物理的音声) を意味づけする「認知」の観点から音節の「重さ」は無視できないのである。ラテンアクセント規則¹⁷⁾ では語末から二番目の重音節に強勢が置かれるが、強勢位置マーカーとして重音節が選ばれるのはそれが物理的に軽音節よりも長くて目立つ (つまり、「重い」) からである。ただし、どのような尾子音が一モーラに換算されるかは言語にも依る。窪菌 (2002) によると、CVVやCVCが二モーラの長さを持つかどうかは、基本的に、VVやVCの部分 (rime) の聞こえ度の度合いと関係するという。つまり、CVCよりCVVが

16) 単位としては点であるものも、具現されると物理的大きさをもつものとなるということである。

17) 脚注3を参照のこと。

二モーラと捉えられやすく、同じCVCでも尾子音が流音 (/l, r/) か鼻音 (/m, n/) の場合の方が阻害音の場合よりも二モーラと捉えられやすいということである。このように、規則において言語間で適用上の確実性が異なるのは、ある意味でもっともなことである。元々、子音と二モーラ母音の第二要素は認識上「点」なのであるから。

英語の尾子音は、最大四個まで連続することができる (Hammond 1999) が、四つ目は必ず形態音素である¹⁸⁾ ので、ここでは本質的に三個とみなして考えることにする。母音の種類と尾子音数の関係には共起制限のあることが分かっている (窪蘭・本間 2002)。一モーラ母音の後には多くの種類のCCが後続するが、二モーラ母音 (長母音・二重母音) の後には極めて限られた種類のCCしか後続しない (例 /st/, /nt/: *beast*, *heist*, *faint*, *pint* /*paint*/)。この事実は、英語にはVVCCの四モーラ音節構造を許容しない規則が存在することを意味する (最大性制約)。実際、英語のアクセント規則に関わる音節の種類をみても、軽いか重いかを二値的に区別するだけで十分であるので、音節量が四モーラとなるような超重音節以上の音節 (VCCC) を想定する必要はないのである。

7. 音節の内部構造

英語音節の内部構造 (syllable-internal structure) に関しては、現在、図5に示すような右枝分かれのものが提案されている (Selkirk 1982, Davenport and Hannahs 1998, Yavas 2006, etc.)。

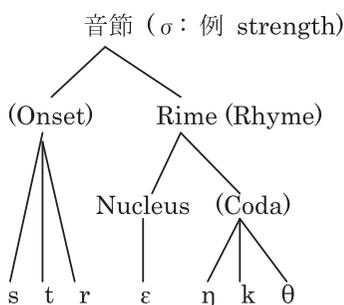


図5. 英語の音節構造

つまり英語の音節は、オンセット (頭子音) とライム (Rime/Rhyme) から成り、そしてライムは核 (Nucleus/Peak) とコーダ (尾子音) から成ると考えられているわけである。尚、オンセットとコーダは、要素としては任意の存在である。この階層構造に基づく知見は、英語においてオンセットとライムが構造上要素としての結びつきが弱いことを示唆する様々な音声現象、例えば韻文化 (頭韻と脚韻)、言葉遊び (e.g. Geta, Pig Latin)、混成

18) 複数マーカー /s, z/, 序数詞マーカー /θ, ð/

語の作られ方 (e.g. breakfast + lunch → brunch)、言い間違い (スプーナリズムや吃音) などの観察結果に基づくものである (窪蘭・太田 1998)。

上述のような、音節構造の要素にライムを普遍的な存在として認める考え方は 1970 年代後半から 1980 年代にかけて支持され (Harris 1983, Treiman & Kessler 1996)、日本語の音節構造も同様の仮定下で分析された (Abe 1987)。一方、後には、英語の音節もその軽重が核とコーダの要素で決まるので、図 6 に示すように、長さを計る中間構成素としてモーラ の概念を組み込んだものも登場する (Hayes 1989)。オンセットは音節の重さに関わらないので、構造上、「音節」に直接支配されているのが特徴である。

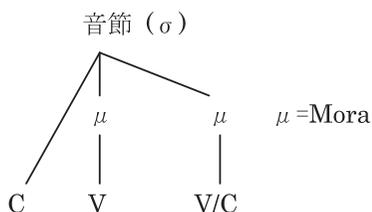


図 6. Mora を組み込んだ英語の音節構造

一方、上述のモーラ の概念を日本語の音節構造分析に応用したものとして窪蘭 (1989) がある。図 7 に示したものがそれである。ここで特筆すべきは、CV とコーダ (C/V) 間に接点が認められ、全体として左枝分かれ構造をしている点である。尚、図 8 に示すような、英語の音節にライムを認めない構造を支持する研究者もいる (Pierrehumbert & Nair 1995)。

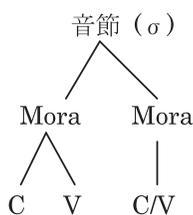


図 7. 日本語の音節構造

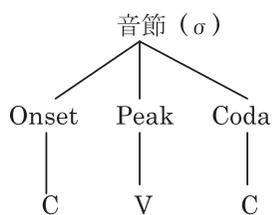


図 8. ライムのない音節構造

要するに、音節の内部構造に関しては今のところ様々な提案があり、研究者により一定しないというのが実情である。では、本稿で先に紹介した幾何学的点・線 の概念に基づく新音節理論に従えば、どのようなタイプの構造が最も支持されるであろうか。本節の最後に以下で考えたいのはこの点である。

すでに上で述べたように、音節とは、長さ計測のために置かれた二点 (起点と着点) で挟まれる一つの連続する「線」のことで、基本的にはリズムの単位である。しかし発話者

が発出した一音節を聴者が同様に一音節（一線）として認知するためには音響学的キューが必要で、これが聞こえ度を基にして成り立つ音素配列上の規則である。よって、音節の定義には、この「音節内で聞こえの山が一つできる」という音素配列論も欠かせない。

さて、モーラ概念を音節の構造分析に導入すると、日本語の音節に関しては図7に示された構造が最も適当であるということになる。日本語の重音節は二つのモーラから成るので、正確に言えば二つの線から成る音節とも言えるが、全体の長さは一セットの起点と着点で計測されており、聞こえの変化においても山が一つしかできない。これが図7の構造を支持する理由である。

一方、英語の音節の場合は、説明がもっと複雑である。Ⅲ.3節で言及したように、音長の計測に一セットの起点と着点を使用されるのはフットのレベルにおいてであり、その構成要素である音節の計測には中点が関与する（フットが一音節語から成る場合を除く）。図9は英語の音節内部構造を示したものである。英語においては、フット中に複数存在できる音節の長さは中点が使われることで分割的に計られる訳だが、聞こえの山が音節数だけできるという点で上述の日本語重音節の場合とは異なる。よって、コーダはオンセットと核を支配するモーラに支配される。理論上、音節という線の末端で点として存在していると考えられるからである。とはいえ、聴者にとっては、開音節と閉音節は長さにおいて同じではない。コーダの物理的長さは音節長を認識する上で0とはみなしにくいからである。だからこそ、英語は音節の軽重に敏感なストレス言語なのである。下図に示したSponge Mora（海綿モーラ）とは、筆者による造語で、音韻論的には長さが0でも、音声学的（物理的）には音節全体の長さに関与するという意味の単位である。海綿はとても軽い水を含むとたちまち重くなるのでこの名を借用したのだが、要は、紙面に表した点のごとく、具現された点は物理的に幅をもってしまっているということが言いたいのである。

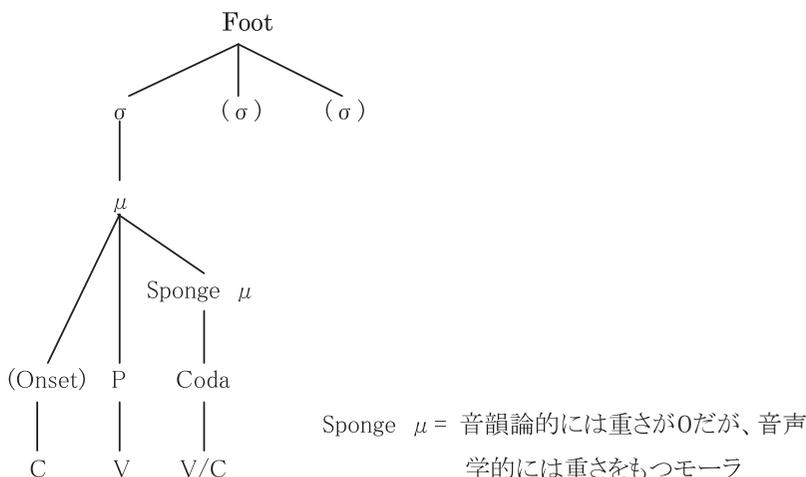


図9. 英語の音節内部構造

本節の結論として英語の音節にもライムは存在しないということになるが、これを支持するには次の質問に答える必要がある。先に紹介した、ライムの存在を示唆するとされる様々な現象（押韻、言い間違い、言葉遊びなど）をどのように捉え直せばよいのかという問いである。

押韻に関しては、よく考察してみると、少なくともライムの存在を確約しているようには思われない。英語になぜ頭韻 (alliteration: e.g. Mickey Mouse) が存在するのかと言えば、最大の理由は、日本語などと比べ頭子音部に起こりうる音の組み合わせの数が多いことにある。例えば、同音の組み合わせが 10 種類の音の中から選ばれる場合と 50 種類の音の中から選ばれる場合とでは、組み合わせにおける希少度は異なり、これが高いほど聞き手は頭韻の存在（意図的な音の選定）に気づきやすくなる (Otaka 2006 p.89)。ライムの韻 (rhyme: e.g. Seeing is believing) は、頭韻とは異なる場所（母音＋コーダ）で起こるが、これが存在できるのも母音と尾子音の組み合わせが希少度において高いからである。英語の押韻には、その他に母音韻 (assonance: e.g. shake and hate) や子音韻 (consonance: e.g. rabies and robbers)、パラライム (pararhyme: e.g. tell and tall) なども存在し、要は、これらによって英語音節の内部構造がオンセット、核、コーダの三要素から成ることが示唆されるのである。

8. 分節

ここで取り上げる「分節」とは、主に行末で起こる正書法上の決まりのことではなく、発話者もしくは聴者が多音節語を分解するときの規則のことである。もっと分かりやすく言えば、分節するとは、多音節から成る音群を最小のリズム単位に分けて生成もしくは知覚する知的営みのことである。尚、一音節語は分節できない。音節長を意識する（計る）ために使われる起点と着点が一つずつしか使われていないからである。

先述の通り、日本語はモーラ拍の言語なので、通常は、(C)V の開音節が分節の際の単位となる。しかし、それ以外にも長音節の (C)VV や (C)VN もある¹⁹⁾。この種の音節は、モーラに着目すれば (C)V.V, (C)V.N と分割が可能だが、前者の場合は分節ではない。全体がその最初と最後にそれぞれ起点と着点をもつ一音節で、分割リズムは関与しないからである。言い換えると、CVV は独立した二モーラ単音節である。よって「東京」は /toʊ/ と /kjoo/ に分節されることになる。一方、(C)VN の場合は、採る音節の定義によって全体で一音節とも見なせるし、(C)V と N（成節子音）に分節することも可能である。前者の解釈は付加リズムにおける音節の定義に基づくわけだが、音節とは「その出だしと括り部に独立した計測アクセント（起点と着点）をもち、核（母音）を中心に聞こえの山が一つ生じる音群」のことである。ここで「独立した計測アクセント」とは、起点と

19) CVQ に関しては、本稿の脚注 6 を参照のこと。

着点を兼ねる中点ではないものことである。よって例えば「簡単」は、/kaN/ と /taN/ に分節される。一方、後者の解釈は分割リズムにおける音節の定義に基づき、音節とは「出だしと括り部に計測アクセントが置かれ、聞こえの山が一つだけ生じる音群」のことである。(C)VN は先の図 2 で示した構造をもち、(C)V と N の長さは分割リズムに則って計測される。この場合、N は成節子音と見なされる²⁰⁾。

英語の分節法はどうであろうか。例えば Japán は、最大頭子音原理 (Maximal onset principle) に基づき Japán と分節される。しかしその形容詞形 Jàpanése では、強勢位置の変動により Jàp.a.nése と分節されるのが普通である (窪 蘭 2002)。英語において強勢音節は必ず尾子音を従えるからである。これはこれで正しい分節法なのだが、もっと正確に言えば、最初の音節 Jap の尾子音 /p/ は先行音節の着点でもあり、後続音節 (/ə/) の頭子音でもあるわけである。このように、分割リズムによって存在している音節を、ポーズを使って物理的に分節することは本来不可能なことなのである。

9. 音節とプロミネンス

プロミネンスとは音節間の聴覚的な相対的際立ちのことで、「卓立」とも呼ばれる。ある音節を他よりも際立たせるには、いくつかの方法がある。通常は、意図的に他の音節よりも音量を上げる (loudness)、高くする (pitch)、長くする (duration)²¹⁾、などである。アクセントは、この卓立を利用して成り立つ音韻規則である。例えば、英語のアクセントでは主に強と弱 (強勢・無強勢) が利用され、日本語のアクセントでは高と低が利用される。換言すると、英語でも日本語でも、強さと高さの単位はどちらも共に音節を単位として付与されるということである (prominence theory of syllable production)。

確かに日本語のアクセントにおいては高さの単位「高」と「低」は音節ごとに与えられるが、中国語などの声調言語では音節内でのピッチ変動は可能である。よって、必ずしも音節ごとに一定の高さが与えられる訳ではないことが分かる。イントネーションにおけるピッチ変動を考えてみれば明らかなおと、ピッチ変動は複数音節に跨る場合も多いのである。ピッチが超分節音素に分類される理由がここにある。

では、強勢付与と音節の関係はどうであろうか。Jones (1950) によれば、ストレスは音響学的に次の四種に分類できる。

1. Level (平板型)
2. Crescendo (次第に強くなる型)
3. Diminuendo (次第に弱くなる型)
4. Crescendo-diminuendo (ピーク形成型)

20) しかし通常は、CVN が全体で聞こえの山を一つだけつくることにより一音節と見なされる場合が多い。

21) 重い音節を使うこともこれに含まれる。

音の高低変化が声帯の緊張と弛緩によって調整されるのに対し、音の強弱は1秒間に肺から押し出される呼気量（もしくは呼気の流れのスピード）で決まる。ストレスの正体は呼気量の増大なのである。この量が増せば共鳴音を作るための声帯振動は激化し、阻害音を作る際の摩擦も大きくなるという訳である。また、このとき音節内では核をピークとして聞こえの山が形成され、聴者が音節を知覚する際の拠り所となる。音節全体が同じ呼気圧（エネルギー）の下で生成されても、共鳴現象を伴う母音の聞こえはその前後の子音よりも大きいからである。

通常の発話において日本語にストレスは適用されないので、上の四種で言えば、1の型が当てはまるであろう。肺を鍛冶屋のふいごに例えれば、この平板型での気圧は常時ほぼ一定といえる。一方、英語の場合はどうか。二重母音において第二母音が第一母音よりも弱体化することや帯気音 [h] が音節頭でだけ起こることなどを斟酌すると、3の型が当てはまる可能性が高い²²⁾。しかしそうではあっても、音節ごとにふいご（肺）に加えられる圧力が毎回変わる訳ではない。例えば 'breath.less.ly' という三音節語を生成するとき、Diminuendo の型はフット全体に適用される可能性が高い。この方が労力上経済的だからである。

IV 音節の新定義に基づく二重母音の考察

音節の定義が済めば、二重母音の定義も自ずと決まることになる。Ⅲ.2節で見た通り、これまでに多くの研究者が「二重母音とは一音節内に収まる（ひとまとまりの、もしくは境い目のない）二種の母音の連続体」のような表現で定義しており、問題は、例外なく音節に関する曖昧模糊な説明にあった。本稿では、要するに、二重母音とは第一要素が音節核として線概念を帯びており、第二要素が点概念を帯びている二種の母音から成る音群のことである。例えばイギリス英語における三重母音の場合（例：'fire' /fɪə/）、核となる母音 /a/ のみが線概念を持ち、第二と第三の要素 /ɪ/ と /ə/ は点として音節中に存在していると見なしうるわけである。また、上の定義により、第一要素がわたり音の音節（例：/ju/, /wu/）は、たとえプロミネンス（ストレス）が第一要素の上から置かれたとしても、二重母音ではないということになる。

英語の二重母音に関して言えば、短母音が11種あるので、理論的にその可能な組み合わせの数は110種類あることになる（ $11 \times 10 = 110$ ）。しかし本稿の冒頭で言及したように、現在、アメリカ英語の場合、そのうち10種のみが二重母音と考えられている。これらを再び挙げると、[aɪ], [eɪ], [ɔɪ], [aʊ], [oʊ], [ɑə], [ɪə], [ɛə], [ɔə], [ʊə] の10種であるが、この内最後の四種が聞こえ度の原則に違反した母音連続となっている。第

22) ただし、いくら Diminuendo の型を取るといっても、頭子音が閉鎖音である音節の場合を除けば、呼気圧が音節頭で常に最大になるというのは考えにくい。ちなみに、日本語のアクセントにおけるピッチ変動でも同じことが言える。音韻論的には（意識の上では）梯子状に変動しても、物理的には比較的滑らかに変動する。声帯振動を低から高へ一瞬にして変えることは物理的に不可能だからである。

一要素が第二要素よりも聞こえの小さいものもあれば ([ɪə], [ʊə])、同程度の組み合わせのものもある ([ɛə], [ɔə])。これは何故なのだろうか。その理由は、先に述べた音節の本質にある。つまり、発話者側に立つと、音節の核音に線概念を付与する限り、第一要素と第二要素はどのような組合せでもよいからである。ただし、聴者側に立つと音群を一つのリズム単位として聴き取るために音響学的な手がかり（聞こえ度）による助けが必要なので、[ɪə], [ʊə] と [ɛə], [ɔə] はあまり相応しい二重母音ではないということになる。では、何故 [ɪə], [ʊə] のような理論に反する二重母音が実際には存在するのだろうか。現時点で筆者が考えている答えは、シュワーのもつ生成上の特徴に起因するというものである。シュワーは中舌母音に分類されるとはいえ、舌を不活動（休憩中）の状態では唇をほとんど開けずに平唇のまま発される。つまり、声道の末端部における調音上の抵抗度 (degree of resistance) は高いままなのである。よってこの母音の聞こえは、高舌母音と大差ないものと考えられる。英語において二重母音 [ɪə], [ʊə] や三重母音 /aɪə/, /aʊə/ が可能なのはこのためであろう。言い方を換えれば、音節末に置かれた曖昧母音シュワーは、英語を母語とする聴者に点として把捉可能な比較的短くて弱い音だということである。

V 英語の二重母音第二要素は単母音かわたり音か

英語の二重母音の第二要素には、通常の単母音記号が使われる場合もあれば、/j/ もしくは /w/ のわたり音（半母音）が使われる場合もある。簡易表記（例：/ai/, /au/）か精密表記（例：/aɪ/, /aʊ/）かに基づいて規則的に使い分けられるときもある。わたり音を使う表記（例：/aj/, /ej/, /aw/, /ow/）よりも単母音を使う表記を好む研究者は、二重母音の第二要素は舌の位置においてわたり音と異なると考えるためである²³⁾。しかし一方で、わたり音を使う表記の方が全体を一つの音節として表現できるので優れていると考える研究者もいる (Hudson 2000, p.30)。では、実際、どちらが支持されるべきかといえ、後者であろう。本稿で示した音節の定義に即しているからである。この考えを採用すると、日・英語間には、音節構造に関して著しい違いがあるということになる。つまり、日本語の音節の着点は必ずしも子音が果たすものではないが²⁴⁾、英語の強勢音節においては必ず子音に依らなくてはならないということである²⁵⁾。着点を子音で示すと音節の区切りが認識されやすくなるからである。

23) つまり、わたり音の場合ほど能動調音器官と受動調音器官による狭窄度が大きくない（高舌化しない）ということである。多くの研究者が指摘している通り、二重母音では母音の質が徐々に変化するが、第二母音の質に完全にたどり着くというわけではない（斎藤 2003）。

24) 理由は、勿論、等時性を帯びるモーラの働きがあるためである。

25) 中点機能を果たす尾子音も含まれる。

VI 日本語に二重母音はあるか

日本語に二重母音はあるか否かの問いに対しては、過去においては否定的な研究者が多かったが（中野 1972, 小泉 1996, 杉藤 1996, 御園・平坂 2013）、現在では肯定的な研究者が多いように思われる。ちなみに、日本語の母音に例外的に高舌母音 /i/, /u/ で終わる二重母音が含まれるのを最初に指摘したのはエドワーズ（1969）であった。そしてその十年後に、『国語学辞典』（時枝編 1979, p.719）の中で柴田が二重母音について解説し、[ei], [ai], [oi], [ui] の四種を二重母音として認めている。ただしその根拠となる説明はさほど決定力をもつものではない²⁶⁾。ともあれ、すでに指摘したように、研究者を問わず、二重母音と連母音の区別はかなり不明瞭なままである。事実、『新版日本語教育辞典』（2005, p.12）における二重母音の記述は、下に示す通り、この不明瞭さを認めたのも同然のものとなっている。丁寧な発音とそうでないときの発音にどんな違いがあるかについて一切言及がないからである。

〔二重母音と連母音〕— 自然な発話では [ai, oi, ui, ae] といった二重母音が観察される（「はい」「おそい」「ねむい」「かんがえる」）。二重母音もそれぞれ一つの二重母音音素の実現とも考えるが、日本語の場合は丁寧に発音すると母音の連続（連母音）が現れる（[ai] など）という事実や音素分析における経済の減速などから短母音音素の連続とみることができる。（『新版日本語教育辞典』2005, p.12）

斎藤（2003）における日本語二重母音の記述も上記のものと大差ない。丁寧な発音かそうでないか、あるいはまたアクセントによって²⁷⁾ 二重母音か連母音かが決まり、常に二重母音である英語の ‘cow’ ([kaʊ]) や中国語の「苦」([kaʊ]) とは異なる点とある（p.7）。また、英語と日本語における二重母音の違いについては次の通り説明している。

英語などにおいては二重母音の終わりの部分は到達点ではなく到達目標ととらえられるので、開始点の違いによって到達点が異なる。例えば、‘high’ [haɪ] の終わりの方が ‘hay’ [heɪ] の終わりより広い。日本語では途中でピッチが上昇する場合は二重母音になりにくい。英語の二重母音は出だしの部分の音色を比較的長く保った上で到達目標の音色に移っていく。それに対して、日本語の場合はすぐに目指すところに移動を開始する。」（斎藤 2003, p.7）

上述のような日・英語間の二重母音の違いは何に起因するのだろうか。私見では、その音

26) 東京語の「営利」と「絵入り」の中のエイの発音を比べると、後者では、二つの母音の間に強さの弱まりがあるのに対し、前者にはそれがなければ「エ」に比べて「イ」が弱い（p.719）、というものである。

27) 日本語では途中でピッチが上昇する場合は二重母音になりにくい。

響学的な違いは二重母音の第二要素が音韻論的に点として解釈されるのか線として解釈されるのかに因っている。

では、結論として、前節で試みた二重母音の定義に基づけば、日本語に二重母音はあると言えるのだろうか。「二重母音では、第一要素は音節核として線概念を帯びており、第二要素は子音同様の点概念を帯びている」と仮定すると、日本語の二重母音（例えば [ai] [au] など）は厳密には二重母音には相当しない。第二要素が線概念を帯びており、全体としては分割リズムで二線が連結された状態にあるからである。第二要素が点として存在する英語の二重母音とは構造が違うのである。では、連母音に相当するかといえば、そうでもない。明らかにリズム構造において連母音とも異なるからである（図 11 参照）。

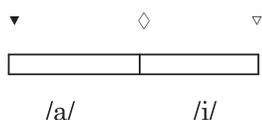


図 10. 日本語の二重母音 /ai/ の内部構造

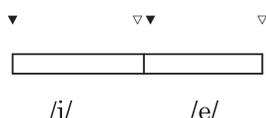


図 11. 日本語の連母音 /ie/ の内部構造

図 11 のような連母音の場合（例：知恵 /tie/）、聞こえの上では全体がピーク一つの山になるが、一音節とは見なせないであろう。母音 /i/ と /e/ は、付加リズムで連結されている可能性の方が高いからである。このように、日本語の /ai/ は、英語の二重母音とは構造が異なるとはいえ、連母音とも異なるのであるから、広い意味で（つまり、音節を生成・知覚のための単位（まとまり：^{おし}節）とする見方に立つと）、二重母音と見なすこともできなくはない。音群の出だしと括り部に音長計測のための起点と着点が意識され、英語の二重母音の場合と同様だからである。

結論として、日本語にも二重母音は存在する、ということになる。ただ、日本語の二重母音は、発話テンポやピッチ変化の影響で連母音として具現されやすいのも事実である。このような条件下では、発話者は分割リズムではなく付加リズムを採用するからである。言い換えると、内在していた中点が起点と着点に分かれて音長計測の機能を果たすということである。

VII 結語

本稿では、重母音（特に二重母音）の本質を音長計測メカニズムの観点から詳しく考察

し、連母音（母音連鎖）との違いを究明した。これまでの研究では、二重母音を定義する際に必ず援用された音節という概念がそもそも明確ではなかったため、二重母音を考える上での障害となってきた。よって本稿では、まず音節とは何かについて一応の新定義を試みてから二重母音の考察に入った。結果、英語と日本語の二重母音には大きな違いがあることが分かった。前者の場合、その第二要素（V2）は幾何学でいうところの「点」として存在しているが（つまり、母音の生成・知覚に際して音長計測上点として処理される）、後者においては「線」として存在しているということである。言い方を換えれば、その際に使用される計測アクセント（起点と着点）の数と置かれ方が双方では異なるということである。

音節は、認知メカニズム上、幾何学でいう線概念を帯びた存在（リズム生成上のチャック）である。構造的に中心となる核がそもそも初めから線概念を帯びた単位として存在しているからである。そしてその前後に付く子音は、幅概念をもたない点として存在する。ゆえに、音群を生成・知覚する際に、全体として一つの線と解釈されうるわけである。よって、子音は一音節内に何個使われようが音節の重さには関与しない。ただしこれは、あくまで理論上の話しであって、実際には物理量をもつために頭子音と尾子音の数の上限は言語によって決まっている。そして、尾子音は音節の重さに関与する。音節中の子音は長さがゼロとして認識されはしても、具現されたときの長さでは核の前と後で異なる。つまり核の前と後ろの子音は、物理的音長という観点で一種の相補分布の関係にあるのである。尚、頭子音と尾子音の数の上限が言語によって決まっているのは、子音の数が多すぎると音節構造が複雑になり、一リズム単位として発音しにくいし、またそのように認知もしにくいからである。そして後者の理由（つまり尾子音はなぜ音節量に関わるのか）については、音韻を音声に具現するときに調音器官の動くスピードが音節末に向かって遅くなるために尾子音は頭子音よりも物理的に長くなりやすいからである。結果、重音節が強勢位置マーカーとして選ばれやすくなるというわけである。

最後に、音節の本質は時間次元上で規定されるリズム単位としての機能にあるのだが、実際には強さ次元における規則（聞こえ度）も利用される。そうすることで、聴者は話者が意図したものと同一リズムを聴取できるようになるからである。

参考文献

- Abe, Y. (1987) Metrical structure and compounds in Japanese, In *Issues in Japanese linguistics*, ed. Takashi Imai and Mamoru Saito, 5-52. Foris.
- Catford, J. C. (2001) *A Practical Introduction to Phonetics* (2nd ed.), Oxford University Press.
- Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, Harper & Row.
- Clark, J & C. Yallop (1995) *An Introduction to Phonetics and Phonology*, 2nd ed., Blackwell.
- Collins, B. and I. Mees (2013) *Practical Phonetics and Phonology* (3rd ed.), Routledge.
- Cruttenden, A. (2014) *Gimson's Pronunciation of English* (8th ed.), Routledge.

- Davenport, M. and Hannahs, S. J. (1998) *Introducing Phonetics and Phonology*, Arnold.
- エドワーズ (Edwards), E. R. : 高松義雄訳 (1969) 『日本語の音声学的研究』, 恒星社厚生閣
- Giegerich, H. J. (1992) *English Phonology*, Cambridge University Press.
- Harris, J. W. (1983) *Syllable structure and stress in Spanish: A nonlinear analysis*. MIT Press.
- Jakobson, R, C. G. M. Fant, and M. Halle (1952) *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*, MIT Press.
- Jakobson, R. and M. Halle (1956/1984) *Fundamentals of Language*, Mouton.
- Jespersen, O. (1926) *Lehrbuch der Phonetik*, 4th ed., Teubner.
- Jones, D. (1950) *The Phoneme: its Nature and Use*, Cambridge: W. Heffer, 3rd edn, revised and expanded, 1967.
- (1960) *An Outline of English Pronunciation*. 4th ed., Cambridge Univ. Press.
- (2011) *English Pronouncing Dictionary* (18th ed.) Edited by P. Roach, J. Hartman and J. Setter, Cambridge University Press.
- 加曾利実 (2013) 日英の二重母音比較論, 聖学院大学論叢第 26 巻第 1 号, 51-62.
- Kenstowicz, M. (1994) *Phonology in Generative Grammar*. Cambridge, Blackwell.
- Kenyon, J. S. (1950) *American Pronunciation*, 10th ed., George Wahr.
- Kenyon, J. S. and T. A. Knott (1953) *A Pronouncing Dictionary of American English*. Springfield, Merriam-Webster.
- 小泉保 (1996) 『音声学入門』, 大学書林
- Kreidler, C. W. (2004) *The Pronunciation of English: A Course Book* (2nd ed.), Blackwell.
- 窪蘭晴夫 (1999) 『日本語の音声』, 岩波書店
- (2002) 「音節とモーラの機能」『音節とモーラ』, 窪蘭晴夫・本間猛 (著), 研究社
- Ladefoged, P. (2006) *A Course in Phonetics*, 5th ed., Thomson Wadsworth.
- Ladefoged, P. and K. Johnson (2011) *A Course in Phonetics* (6th ed.), Wadsworth.
- Ladefoged, P. and I. Maddieson (1996) *The Sound of the World's Languages*, Blackwell.
- Lodge, K. (2009) *A Critical Introduction to Phonetics*, Continuum.
- 桥矢好弘 (1976) 『英語音声学』, こびあん書房
- McMahon, A. (2002) *An Introduction to English Phonology*. Edinburgh: Edinburgh University Press.
- 御園一夫・平坂文雄 (2008) 二重母音と二母音連続『関東学院大学文学部紀要』第 113 号, 関東学院大学, 87-99.
- 中野一夫 (1972) 『英語母音論』, 学書房出版
- 中島直嗣 (2011) tower/towel, hire/higher の発音について—英語のいわゆる「三重母音」の再考察 *Setsunan Journal of English Education (SJEE)*, 5, 55-68.
- (2012) 英語の平滑化と三重母音の本質, 日本英語音声学会 (EPSJ) 学術論文集『英語音声学』 16, 125-135.
- (2015) 英語のわたり音 /j/ の音韻論的分析, 日本英語音声学会 (EPSJ) 中部支部創立 20 周年記念第 23 回研究大会 2015 年 5 月 30 日 (土)
- 日本語教育学会 [編] (2005) 『新版日本語教育辞典』, 大修館書店
- 小野浩司 (2011) 二重母音とは何か, 佐賀大学紀要 Vol.15, No.2, 189-197.
- 大高博美 (1987) 日本語の音節構造とリズム, 『月刊言語』, Vol.16, No.6, 82-92, 大修館書店
- (1988) 音節とは何か, モーラとは何か, 『月刊言語』 Vol.17, No.3, 60-68, 大修館書店

- (2012) 英語における Spondee リズムの特徴, 『21 世紀英語研究の諸相』 274-289, 開拓社
- Otaka, H. (2006) *Phonetics and Phonology of Moras, Feet and Geminate Consonants in Japanese*, University Press of America.
- (2013) The phonological substance of Q and the bimoraic foot in Japanese. *Kwansei Gakuin University Humanities Review* Vol.18, Kwansei Gakuin University.
- Padgett, J. (2008) Gildes, vowels, and features. *Lingua* 118, 1937-1955.
- Paradis, C. and D. LaCharité (2014) Loanword adaptation: from lessons learned to findings. In Goldsmith, J. A., J. Riggle and A. C. L. Yu (eds.) *The Handbook of Phonological Theory* (2nd ed.) 751-778, Wiley-Blackwell.
- Pierrehumbert, J. and Nair, R. (1995) Word games and syllable structure, *Language and Speech* 38, 77-114.
- Pullum, G. K. and W. A. Ladusaw (1996) *Phonetic Symbol Guide* (2nd ed.), University of Chicago Press.
- Roach, P. (2009) *English Phonetics and Phonology: A Practical Course* (4th ed.), Cambridge University Press.
- Roca, I and W. Johnson (1999) *A Course in Phonology*, Blackwell.
- Sachs, C. (1953) *Rhythm and tempo: A study of music history*, W. W. Norton.
- 斎藤純男 (2003) 「現代日本語の音声—分節音と音声記号」『朝倉日本語講座 3 音声・音韻』(北原保雄 [監修], 上野善道 [編]) 第 1 章, 朝倉書店
- Schubiger, M. (1970) *Einführung in die Phonetik*, 2nd ed. Berlin: Walter de Gruyter. [小泉保訳『音声学入門』, 大修館書店 1973]
- Selkirk, E. (1982) The syllable. In the structure of phonological representations (part II). Ed. Harry van der Hulst and Norval Smith, 337-383. Forris.
- (1984) On the major class features and syllable theory. In Aronoff, M. and R. T. Oehrle (eds.) *Language Sound Structures*, 107-136, MIT Press.
- 志賀浩二 (1992) 『数の世界』, 岩波書店
- 杉藤美代子 (1996) 『日本人の英語』, 和泉書院
- 竹林滋 (1996) 『英語音声学』, 研究社
- 時枝誠記 (編集委員長) 国語学会編 (1979) 『国語学辞典』東京堂
- 鳥居次好・兼子尚道 (1969) 『英語発音の指導』, 大修館書店
- Trask, R. L. (1996) *A Dictionary of Phonetics and Phonology*, Routledge.
- Treiman, R. and Kessler, B. (1995) In defense of an onset-rime syllable structure for English, *Language and Speech* 38, 127-142.
- Wells, J. C. (2008) *Longman Pronunciation Dictionary* (3rd ed.), Pearson.

The Phonological Differences between Diphthongs and Compound Vowels in English and Japanese

Hiromi OTAKA

Abstract

The purpose of this paper is to explore the phonological difference between diphthongs and compound vowels in English and Japanese. Diphthongs in American English for example are usually thought to be vowels like [aɪ] , [eɪ] , [ɔɪ] , [aʊ] , [oʊ] , [ɪə] , [eə] , [ʌə] , [ɔə] , and [ʊə] (Yasui 1992). In Japanese, some compound vowels like [ai] and [au] are also sometimes regarded as diphthongs (Kubozono and Homma 2002).

Diphthongs have been defined in various ways by many phoneticians, but they have not yet been studied enough in order to reach a precise definition. Many linguists consider diphthongs as “a single syllabic nucleus which begins with one vowel quality and changes more or less smoothly to a second vowel quality” (Trask 1996). However, this definition is unclear because it cannot successfully differentiate diphthongs from compound vowels. For example, Japanese compound vowels such as /ie/ (‘house’) and /ei/ (‘thornback fish’) also appear to be connected smoothly in terms of quality, judging by F1 and F2 on a sonogram.

Thus, researchers have come to also consider sonority (i.e. intrinsic loudness) in order to make the definition of diphthongs more precise, claiming that “one of the two vocalic elements is usually more prominent in sonority than the other, as in /ju/ (V2 > V1) and /ai/ (V1 > V2)” (ibid) . However, once again, this seems to not be a precise definition because diphthongs like /ju/ (V2 > V1) could also be regarded as CV composed of a glide plus a vowel. In addition, a definition based on the latter case (V1 > V2) seems to not always be able to differentiate diphthongs from compound vowels in Japanese. It is still not clear whether for example /ei/ is monosyllabic or disyllabic. Furthermore, accentuation could also be influential in differentiating diphthongs from compound vowels in Japanese. For example, aside from sonority, the vowel sequence /ai/ in the word *maiku* (‘microphone’) can be regarded as a diphthong because V1 is accented, but its counterpart in another word *taisho* (‘object’) can be regarded as a compound vowel because V1 is not accented in this word.

In this paper, diphthongs are defined based on a new definition of the syllable, which results in a clear differentiation between diphthongs and compound vowels.

The syllable is basically a rhythmic unit composed of one nucleus and a few (or no) consonants before and after it, and diphthongs are vowels occurring within a nucleus. In a geometric sense, a nucleus is phonologically analogous to a line, while consonants are analogous to points, so a diphthong can be regarded as a single line consisting of points in a sequence, of which the second V occupies the last point. The physical duration of a syllable is realized by means of a phonological unit “mora.”

On the other hand, compound vowels are composed of two vowels existing phonologically as two lines in a row. In other words, the second V is regarded as a line, not a point.

The concept of sonority is also useful to define diphthongs more accurately because it can help listeners grasp a sequence of phonemes acoustically as a chunk called a “syllable,” which consists of one peak in intrinsic loudness.