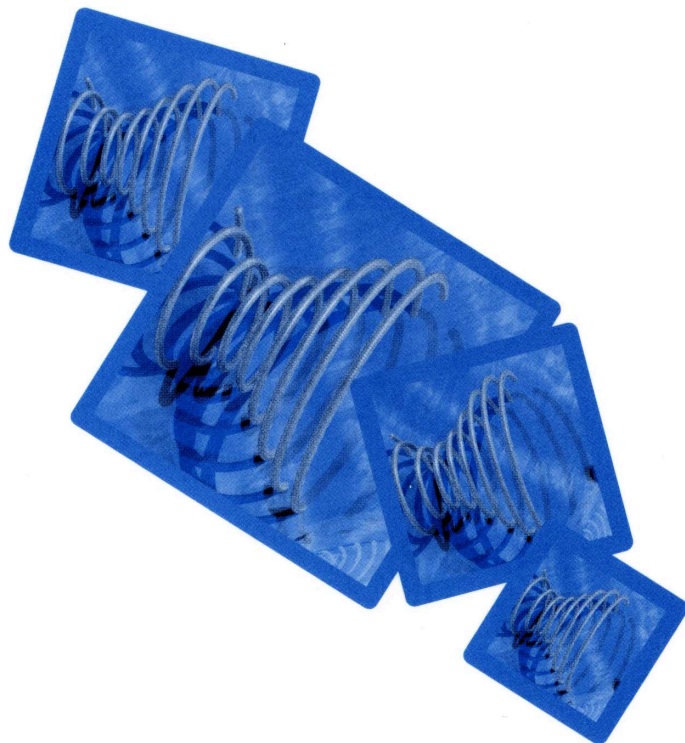


音声工学

板橋 秀一 編著

赤羽 誠 石川 泰 大河内正明
粕谷 英樹 桑原 尚夫 田中 和世
新田 恒雄 矢頭 隆 渡辺 隆夫
共 著



編著者略歴

板橋 秀一 (いたばし・しゅういち)

1970年 東北大学 大学院 工学研究科博士課程 単位取得退学
東北大学 電気通信研究所 助手
1972年 工業技術院 電子技術総合研究所 技官
1974年 同 主任研究官
1977 - 1978年 スtockホルム王立工科大学 客員研究員
1982年 筑波大学 電子・情報工学系 助教授
1987年 同 教授
2004年 筑波大学 大学院システム情報工学研究科 教授
2005年 筑波大学名誉教授
産業技術総合研究所 客員研究員, 国立情報学研究所 特任教授
現在に至る. 工学博士

音声工学

© 板橋秀一 (代表) 2005

2005年2月15日 第1版第1刷発行 【本書の無断転載を禁ず】
2006年2月10日 第1版第2刷発行

編著者 板橋秀一
発行者 森北 肇
発行所 森北出版株式会社
東京都千代田区富士見 1-4-11 (〒102-0071)
電話 03-3265-8341 / FAX 03-3264-8709
<http://www.morikita.co.jp/>
日本書籍出版協会・自然科学書協会・工学書協会 会員
JCLS < (株) 日本著作出版権管理システム委託出版物 >

落丁・乱丁本はお取替いたします

印刷/モリモト印刷・製本/石毛製本

Printed in Japan / ISBN4-627-82811-X

森北出版株式会社

時間軸を含めて考える必要がある。/j/と/w/は、調音の構えとしては母音の/i/と/u/に類似しているが、典型的な発声の場合には調音位置での狭めが強い結果、/j/では F_1 と F_2 の間が幾分広くなり、 F_2 が弱くなる傾向がある。/w/では、 F_2 が/u/よりやや低くなる。しかし、通常の発話中の特性は前後の音韻に影響される。例えば、/aja/と発声したときの/j/に相当する区間の F_1 、 F_2 の特性は/i/より/e/のそれに近い。半母音としての特徴は、調音に基づく直観からは、 F_1 、 F_2 の時間変化パターン、特に両者の位相関係にあると考えられるが、定量的な結果を示して検証した例は少ない[8]。なお、過渡的な部分とは別に、半母音に相当する区間長あるいは後続母音の定常的な区間が短い場合に半母音と知覚される傾向もある。

(2) 破裂音

破裂音には、無声破裂音/p, t, k/と有声破裂音/b, d, g/がある。前者について音響現象を時間軸に沿って観察すると、先行母音からの過渡部、無音部(閉鎖部)、破裂部、気音部、母音への過渡部(有声音区間)と続く。これに対して有声破裂音は破裂の時点で声帯が振動している音であり、典型的には先行母音からの過渡部、声帯のバズ音(閉鎖部)、次に破裂部、母音への過度部が現れる。無声破裂音/ka/と有声破裂音/de/の波形を図2.13に示す。無声破裂音と有声破裂音の差異は、基本的には破裂の瞬間から声帯が振動を始めるまでの時間 voice

2.3.5 子音の性質

子音は、母音のような定常的な性質をもたないものが多く、声道の途中に強い狭め(閉鎖)があつて、音源がこの狭め付近で生成され、その音源の特性も様々である。また、鼻音のように声道に分岐管ができる場合がある。以上のような点で音響理論的扱いが難しく、スペクトルの特徴も解析的な形では明解に記述できない場合が多い。子音の音響的特性は、音源の形態や声道の分岐、声道の閉鎖(狭め)の強さなどに相当する調音様式と、声道の狭めの位置にあたる調音点とによって大きく分類される。ここでは、日本語の主な子音について、それらの識別という観点に重点をおいて、音響分析の特徴を明らかにする。

(1) 半母音

半母音にはヤ行音/j/とワ行音/w/がある。調音位置での声道の狭めが母音よりやや強いが、音響理論的には母音と同様に扱える。ただし、その特徴の記述は

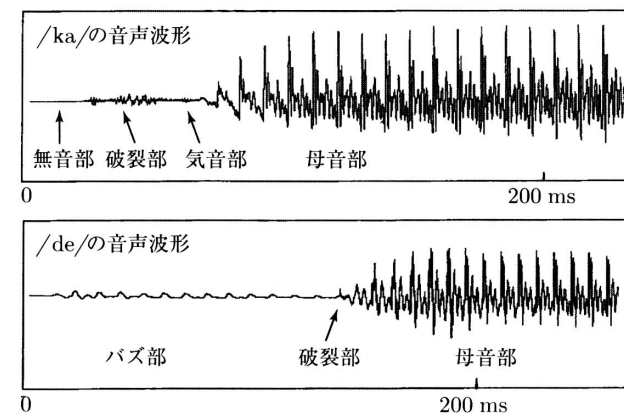


図 2.13 無声破裂音と有声破裂音を含む単音節音声の波形

onset time (VOT) にあり、欧米語では、VOT が 20 ミリ秒付近を境にして、これより長い場合が無声破裂、短い場合が有声破裂とされる。しかし、日本語の通常の発声では必ずしもこの数値は当てはまらず、VOT が 10 ミリ秒前後の(気音部が認められない)無声破裂音も /p/ などに多い。

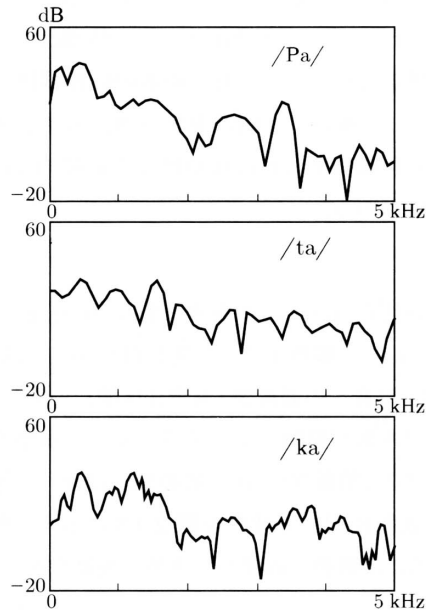


図 2.14 後続母音が /a/ のときの無声破裂音の破裂部のスペクトル

破裂音群内での互いの区別は、調音点の差異に起因するパワースペクトルパターンの特性に依存する。ただし、これらは前後の音韻環境に影響を受ける。声道の共振モードを考慮すると調音点の違いが後続母音への渡りの部分にも現れると考えられるが、音響分析によれば無声破裂音では特徴の主要な差異は破裂部分にある。図 2.14 にこの破裂部分のパワースペクトルのサンプル例を示す。無声破裂音のスペクトル概形の特徴をまとめると次のようになる [9][10]。

- ① /k/ のスペクトルの特徴は、スペクトルが集約的であること、すなわち、後続する母音の F_2 に接続する付近の帯域に強いスペクトルのピークが現れる。

- ② /t/ のスペクトルは拡散的であり、特に強いピークはもたない。
- ③ /p/ は /t/、/k/ に比べ相対的に低い帯域の成分が強い。

また、無声音区間長について観ると、/p/、/t/ が /k/ に比較して概ね短い。無声破裂音を精度良く類別するには、後述する有声破裂音の場合のように前後の母音に依存して特徴を記述する必要がある。

有声破裂音 /b, d, g/ の特性は、それらの調音点が /p, t, k/ の場合に類似しているので類別的特徴も似ている。しかし、自動的に類別するのは無声破裂音の場合より幾分困難になる。有声破裂音は、 F_2 、 F_3 の後続母音への遷移パターンに差異があるとされるが、前後の音韻環境を考慮する必要がある。図 2.15 に後続母音が /a/ の場合のスペクトログラムのサンプル例を示す。後続母音別に、その特徴を述べると次のようになる [11]。

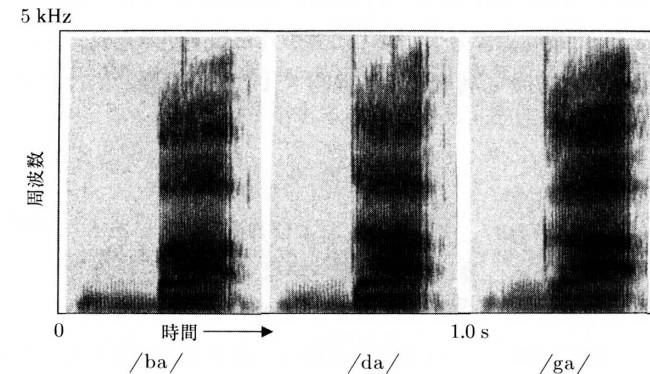


図 2.15 有声破裂音を含む単音節 /ba/, /da/, /ga/ のスペクトル

- ④ 後続母音が /i/ の場合。/g/ の破裂部分からの F_2 の軌跡 (F_2 ローカスと呼ぶ) は、/i/ の F_2 よりも高く、定常的な部分が数十ミリ秒ある。これに対して /b/ の F_2 は低い方から始まり、急な変化をする。/d/ は特性的には両者の中間であるが、日本語共通語では口蓋化して [dzi] (破裂音「ジ」) となる。
- ⑤ 後続母音が /e/ の場合。/g/ に関しては /i/ の場合と同様、ただし、 F_2 がやや低い。/b/ では F_2 ローカスは /e/ の F_2 よりやや低い方から始まる

が、変化量は少ない。/d/は、/b/に比べローカスが幾分高いが、ほぼ同じ傾向で、両者の差異は小さい。

- ⑥ 後続母音が/a/の場合。/g/と/d/は、共に F_2 ローカスが/a/の F_2 より高いが、/d/の方がより高いところから始まり/a/の定常部に向かって下がる。/b/は相対的に F_2 が低い。
- ⑦ 後続母音が/o/の場合。/g/と/b/の F_2 ローカスは、共に、/o/の F_2 とほぼ同位置であるが、/g/の F_2 の方がやや強い。/d/の F_2 ローカスは高く、/o/の定常部に向かって下る。
- ⑧ 後続母音が/u/の場合。/g/と/b/の F_2 については、/o/の場合と同様である。ただし、全体にやや高い。/d/の F_2 ローカスは相対的に高いところが下がるが、日本語共通語の発声では [dzu] ないし [zu] の音になる。

なお、上記のいずれの場合にも、 F_1 は低い周波数から母音に向かって上がる。また、/g/の F_2 は後続母音の F_2 , F_3 に分かれる傾向が見られ、これは調音と声道の共振モードに関する考察からも推測されるが、音響的には必ずしも明確には観測できない。

(3) 鼻子音

鼻子音には、マ行音/m/, ナ行音/n/および鼻濁音/ŋ/がある。他に「ん」に相当する撥音/N/がある。鼻子音の音響現象は、まず、先行母音からの入り渡り、ほぼ定常的な鼻音部 (nasal murmur), 母音への出渡りと続く。ただし、鼻音化の現象 (口蓋帆が開いている状態) 自体は、鼻子音の区間のみに限定的に生起するものではなく後続する母音などの区間に持続することが多い。鼻音部のパワースペクトルは、図 2.16 に示すように、口腔から鼻腔への分岐に起因する反共振により零点ができる [12]。この周波数位置は口蓋帆から口腔の調音点までの長さに関係するため、この長さが短い n/の方が/m/より高くなる。成人男性の声では、/m/では概ね 0.8 kHz 以下、/n/ではそれより高い位置にある。

音響的特徴に基づいて鼻子音群を識別するのは、一般に、かなり困難である。上述したスペクトルの零点を自動的に抽出するのは難しく、通常はスペクトル包絡の概形やホルマント周波数を用いる。合成音を用いた知覚実験の結果などからは、/m/, /n/などの違いは、主に母音への渡り部にあるが、鼻音部にあると考えられる場合もある。/m, n, ŋ /について、渡り部の第 1, 2 ホルマントの

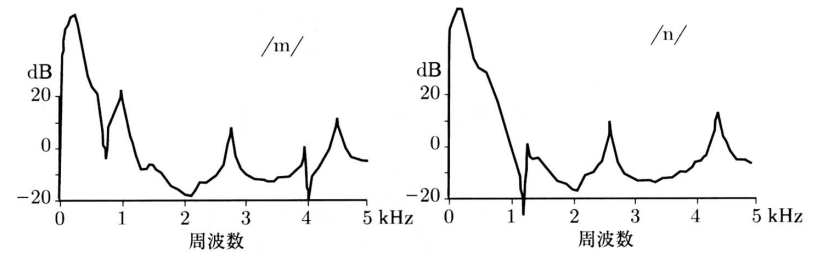


図 2.16 鼻子音 /m/ と /n/ の典型的なスペクトル。声道断面モデルから計算したもの [12]

動きは、(2) で述べた有声破裂音/b, d, g/の場合にほぼ対応している。

/N/は、他の鼻子音に比べ、前後の音韻の影響を受け易く、典型的には後続子音の調音点に依存する。例えば、後続子音が/p/, /b/のときは/m/に近く、/t/, /d/のときは/n/に近い音になる。また、終端では先行母音の特性に鼻音化が重畳されているように見える場合もある。なお、鼻音化母音の特徴は、反共振の影響によって見かけ上、本来の母音ホルマントに加えて新たにピークが生じたように見える特性を示す。

(4) 摩擦音

摩擦音には、無声摩擦音であるサ行音/s/とハ行音/h/, 有声摩擦音のザ行音/z/がある。音響現象的には、調音点に音源があり、定常的で比較的安定している。サ行音の音素/s/には、[s] と後続母音が/i/の場合の [ʃ] がある。後者は日本語の音素表記として 2 音で表す/sj/(シャ行音) とほぼ同じで、/sj/は、音声的には 2 連音というよりは単音 [ʃ] と見てよい。[s] や [ʃ] のパワースペクトルは、概ね、調音点より唇側の共振特性によって表される。その例を図 2.17 に示す。この図から分かるように、3 kHz 以上の高域にパワー成分の現れる周波数が唇から調音点までの長さに概ね反比例するので、両者の差異がこの周波数位置の違いに現れ、[s] の方がこの位置が相対的に高い [13]。

ハ行音の子音/h/は、後続母音が/u/の場合を除くと、後続母音の声道の狭め付近に摩擦音源ができる。このため、後続母音が/a, o/のときはこれらの母音の F_1 , F_2 の中間付近、/i, e/の場合は同じく F_2 , F_3 の中間付近にスペクトルのピークができる。/hu/は、「フ」に対応する音素表記であるが、音響的には

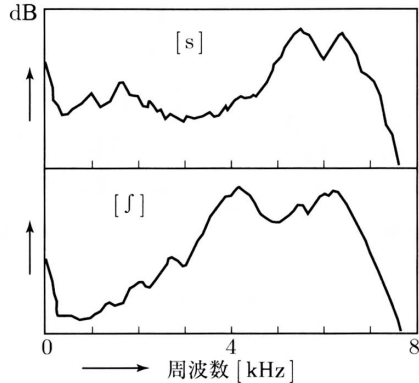


図 2.17 無声摩擦音 [s] と [ʃ] のスペクトル

「ファ」などに対応する表記/fa/の/f/に近く、発声の個人差などによる特性の変動も大きい。

有声摩擦音は、通常、バズ音が先行し、これに摩擦性の音が重畳し、後続母音へと移行する。/s/と同様に後続母音が/i/の場合、口蓋化して「ジ」[dzi]になる。「ジ」は、単独で発声されるときは概略、有声破擦音のカテゴリーに入るが、語中などでは前後の環境に依存する。一般に、/z/では声帯音源も重畳されるので、/s/と同様な高い周波数成分に加えて、500 Hzより低い帯域にパワー成分が現れる。

(5) ラ行音

日本語のラ行音/r/は、音素としての生起頻度は/k/などと共に高いが、その音響的特性の研究は比較的少ない。音声学的には、弾き音 (flapped)、側面音 (lateral)、ふるえ音 (trill) などに分類される。音韻環境や人によって特性が大きく変わり、連続音声の中では音響的变化が小さく検出が難しい場合もある。調音点の類似から音響的特性が/n/、/d/と近く、それらとの類別はかなり困難である。定性的には、舌尖が口蓋に接する閉鎖時間が短く、この閉鎖区間への入り渡り、出渡りのホルマントパターン、パワーディップに特徴を見いだすことができる [14]。