

『日本語話し言葉コーパス』における鼻子音の持続時間

高 田 三 枝 子

1. 本稿の目的

本稿は日本語の音節頭の鼻子音 /m, n/ に注目し、その持続時間の分布を現代日本語の大規模自発音声コーパスである『日本語話し言葉コーパス』(Corpus of Spontaneous Japanese: CSJ) を用いて分析した結果を示すものである。(以下、特に言及しない限り、本稿で言う鼻子音はすべて音節頭のものである。)

既に知られているように、音声的に類似しており同じ記号や同じ素性で表される音声でも言語あるいは変種間で完全に一致するわけではない (Keating 1984)。筆者はこの、言語あるいは変種における音声詳細の実態とその言語内のおよび外的要因の下での分布に興味を持っている。これらの音声詳細は、現時点で弁別に関わるものでなかったとしても、何らかの音声変化を反映している場合もあるからである。分節音の持続時間の分布はそうした音声詳細の基本的情報の一つである。

本研究で鼻子音の持続時間を取り上げるのは、この鼻子音の持続時間が voicing の開始のタイミングに関わるからである。話頭位置における有声破裂音においては voicing の開始は voice onset time (VOT) として測定される。この VOT に関しては、地理的分布および世代的分布に関して非常に興味深い変異が観察される。すなわち、VOT がマイナスになるか否か (voicing lead を伴うか否か) ということに関して、地理的には東北とそれ以外とで大きく異なり (東北はプラス)、世代的には若い世代ほどプラスの値をとることが明らかにされている (高田 2011)。

鼻子音 /m, n/ は調音動作に関して有声破裂音と共通する点が多い。すなわち両者はその鼻音性の有無において異なるが、どちらも口腔内の声道を閉鎖し、その閉鎖区間中に声帯振動が始まり、口腔の閉鎖開放後に母音が後続するという一連の調音動作を行う点で共通する。そうした時間軸上の一連の調音動作の対応において、鼻子音の持続時間 (音響分析の観察上、鼻音マーマーと呼ばれる部分) は口腔閉鎖区間中の声帯振動によるエネルギーが観察できる区間という意味で有声破裂音における voicing lead に対応するものとして考えられ、有声破裂音において観察された地域差・世代差が、鼻子音においても観察される可能性は大いに考えられる。こうした議論に対する準備的な分析として、まずいわゆる共通語の発話における音節頭鼻子音の標準的な持続時間の在り方について、各音環境の条件の下での変異を含め、明らかにしておきたい。すなわち地域的・世代的多様性という言語外的な条件以前に、言語内的な条件の当該分節音の持続時間への影響を明らかにしたい。

本稿で扱う CSJ は、後にも述べるように（第 3 節）「東京語を基盤とした標準的日本語音声である」とみなすことのできる（前川 2010）大量の音声資料に対して分節音ラベリングが施されたコーパスであり、これを分析することで日本語の音声実態の基礎的情報を効率的に把握することができる。

鼻子音を含めた各音素に対応する単音の持続時間については、これまでも日本語におけるモーラの単位の設定の妥当性の議論（Han 1962、Beckman 1982）や、音声合成における時間制御の議論（匂坂・東倉 1984、Campbell 1992、舛田・他 2002）の中で示されている。しかしそれらで使用した音声資料は単語リストや文章の読み上げによるものであり、自発音声によるものではない。前川（2011）が指摘するように、コミュニケーションにおいて発せられる音声特徴には、読み上げ音声では捉えきれない側面があり、自発音声における音声詳細を、これまでの知見に加えて明らかにすることには意義があると考えられる。

本稿では、先行研究の報告に対し、より自然な発話状況の日本語において音声特徴がどのような分布を示すのかという観点から報告を行う。ここで示す結果は、今後も行われるであろう読み上げ音声による研究に対して、より多くの要因の関わる自然な日本語とどの程度の違いが見込まれるかということについても一つの資料を与えることができると考える。

2. 先行研究

本節では、鼻子音の持続時間に関わる要因および先行研究で示された具体的な持続時間の数値について述べる。

2.1. 鼻子音の持続時間に関わる要因

日本語の分節音の持続時間に関わる要因としては、当該音の音種の他に、発話速度、隣接音、当該音の含まれる呼吸段落や句の長さ、当該音の発話単位内での位置（文、呼吸段落、アクセント句（以下、AP）、など）、統語的属性などが指摘されている（板橋 2005：159）。特に子音については、当該音の音種が大きく影響することが知られている（匂坂・東倉 1984、Kaiki & Sagisaka 1992）。

発話速度の影響は子音にも観察されるものの、影響は母音でより大きいとされる。舛田・他（2002）によれば「子音では全体的に縮める方向の伸縮は行われやすいが、伸ばす方向への伸縮は行われにくい」という。隣接音の影響は、時間長補償現象として知られるもので、「モーラをタイミング単位とした、いわゆる音節リズム（syllable timed rhythm）の現われである」（匂坂・東倉 1984：632）とされる。ただし子音は母音ほど隣接音の影響は受けないという（匂坂・東倉 1984、Kaiki & Sagisaka 1992）。また Kaiki & Sagisaka（1992）では、他に、発話単位中での位置の影響については、文内位置の影響や AP 末での伸長などはあまり見られないが、呼吸段落末や AP 頭での伸長があること、呼吸段落や句などの長さの影響は子音に対しては非常に小さいこと、品詞の違いの影響は小さいが見られ、その中で数量詞や固有名詞、重要情報の伝達・トピック表示・接続を示す助詞などは伸長するが、機能語は一般的に短縮する傾向があることを述べている。なお呼吸段落末や AP 頭の影響については、調音法との関係も指摘されており、前者については /r/、後者については /p, t, k/ への影響が大きく、これに比べ鼻子音への影響はあまり大

きくないという。

本稿では以上の先行研究で影響があるとされる要因を中心に分析を行うことにする。すなわち AP の発話速度、語の種類、調音位置、隣接音（ポーズの有無、先行音¹、後続音）、AP 内位置（AP 頭か否か）、といった要因による持続時間の分布を分析する。なお、先行研究では呼気段落末で伸長が起こることも指摘されているが、これについては呼気段落という単位情報を CSJ でどのように抽出するか検討が必要であり、今回は分析対象としなかった。

2.2. 先行研究で示された鼻子音の持続時間

本節では先行研究で使用された音声資料と報告された鼻子音の持続時間について整理して示す。表 1 に先行研究で使用された音声資料の概略を示し、また表 2 に各先行研究で示された /m, n/ の平均持続時間を示す。

表 2 を見ると、/m, n/ 両方の値を示している先行研究においては、全体に /m/ > /n/ となっていることが読み取れる。またこれまでに示された値は先行研究間でばらついているが、これらの研究を通してみると、音声収録において話者の注意が発音に届きにくい条件の音声資料において得られた値ほど短い値を示していると見ることができる。すなわち長文読み上げの音声資料を用いた Hiki et al. (1967) の報告がもっとも短い値を示し、匂坂・東倉 (1984) の「/VCV/ 音韻連鎖」(単独読み上げと思われる) の値はこれの倍近い値を示している。(なお Han (1992) は短文読み上げでありながら長い値を示しているが、これは全資料中から 1 つしか取り上げていないなど資料の扱いに恣意的なところがあり参考にとどめるべきであろう。)

以上に示したように、これまで先行研究で示されてきた鼻子音の持続時間の値は、読み上げ音声によるものであり、本研究ではこれらの先行研究で示された値に対して、自発音声の資料では

表 1 分節音の持続時間の測定値を示した先行研究

論文名	話者数	話者情報	資料収録手法	資料規模	持続時間記載音*
Han (1962)	3	なし	短文読み上げ	約 350 文? (母音約 300、子音約 50 のスペクトログラム)	/m,n,k,N/
Hiki et al. (1967)	1	男性 (アナウンサー)	長文読み上げ (天気予報)	音素数 778、 ポーズ数 51	/a,i,u,e,o,y,w,m,n,ng,r,p,t,k,s, h,b,d,g,z,N,Q/、ポーズ
Beckman (1982)	5	女性 1、男性 4 (標準語話者)	短文読み上げ (有意味語) 読み上げ	1125 語 (75 語 × 15 回)	/a,u,e,o,m,r,t,k,s,b,d,g,z,Q/
匂坂・東倉 (1984)	1	男性 (標準語話者)	(/VCV/ 音連鎖) (無意味語単独) (有意味語単独・短文)	1600 発話 (310 発話 × 4 種)	/a,i,u,e,o,y,w,m,n,r,p,t,k,s,h,b, d,g,z/
Campbell (1992)	1	男性 (アナウンサー)	短文読み上げ	503 文 (母音数 10957、 子音数 9529)	/a,i,u,e,o,R,y,w,m,n,r,t,k,s,h/
舩田・他 (2002)	2	女性 (ナレーター)	短文読み上げ	525 文 (「ATR 音素バランス 503 文」、外来語音韻追加用 セット 22 文)	/a,i,u,e,o,R,y,w,m,n,r,p,t,k,b, d,g,s,z,h,N,Q/、ポーズ

* 異音や音連鎖を除く

どのような結果になるかが注目される。

表2 先行研究で示された鼻子音の持続時間

論文	鼻子音分析数 (m/n)	平均持続時間 (ms)		話頭 (ポーズあり) /話中
		/m/	/n/	
Han (1962)	1-1	90	70	話中のみ
Hiki et al. (1967)	32/30	49	51	話頭および話中
Beckman (1982)	210/0	60 前後*	なし	話中のみ
句坂・東倉 (1984)	記述なし	/VCV/: 94 無意味語: 70 有意意味語: 62 短文: 56	/VCV/: 86 無意味語: 69 有意意味語: 60 短文: 56	話中のみ
Campbell (1992)	763/1119	46	41	話頭および話中
舛田・他 (2002)	記述なし	速度速い: 70 前後* 速度普通: 80 前後* 速度遅い: 95 前後*	速度速い: 50 前後* 速度普通: 60 前後* 速度遅い: 75 前後*	話頭および話中

* 図から読み取った値。

3. 資料

本研究では、先にも述べたように、CSJを使用してその持続時間を分析するが⁸ (CSJの仕様、また音声の転記やラベリングの詳細については前川 (2004)、国立国語研究所 (2006) 等を参照されたい)、その中でも、X-JToBIによる音声ラベリングが施された「コア」(以下、CSJ-Core)の独話(学会講演(以下、APS)と模擬講演(以下、SPS))部分のみを分析対象とする。同分析対象データは前川 (2010) で用いられたものと同じであり、以下前川 (2010) の説明に従って簡単に述べる。

CSJ-CoreのAPSとSPSの話者数は異なりで123名である。APSの話者は大学院生が多く、SPSの話者は「人材派遣会社からの派遣によって年齢と性別をほぼバランスさせた話者群」とされている。出生地は東京都92名、千葉県12名、神奈川県12名、埼玉県3名、その他4名(東京圏での生活歴が長い)となっている。前川 (2010) と同じく、本稿ではこの音声資料を「東京語を基盤とした標準的日本語音声である」とみなす。話者数等の情報を表3に示す。なお表中の「短単位語数」とは例えば「国立国語研究所」であれば「国立|国語|研究|所」のように区切る単位である。

本稿の分析対象である鼻子音の持続時間情報は、CSJ-Coreから/m/、/n/とについて抽出した。また同時に、当該発話の基本情報(ID、話者情報など)、APと短単位(それぞれ、当該、直前、直後)の情報、当該鼻子音とその前後それぞれ4セグメントの情報、直前のポーズの有無、鼻子音を含むモーラの位置(AP中、長単位中、短単位中)についての情報を抽出した(抽出作業は前川喜久雄氏に行っていた)。)

表3 CSJ-Coreの独話資料

	話者数*	短単位語数	時間数**
APS	23/45	218161	14.2
SPS	38/37	225572	15.0

前川 (2010) 表1をもとに作成

* 斜線の左は女性、右は男性

** 発話間のポーズを除去した実質発話時間

4. 分析

先に見た、子音の持続時間に関する先行研究の報告から、本稿では鼻子音の総合的な持続時間の分布(4.1.節)の他、APの発話速度(4.2.節)、語の種類(自立語、付属語、感動詞)(4.3.節)、調音位置(4.4.節)、隣接音(4.5.~4.5.3.節)、AP内の位置による違い(4.6.節)を分析する。なお以降の分析ではポーズを挟まず撥音 /N/ に後続する音節頭の鼻子音はそれ自体の持続時間を /N/ から分けて計測することができないため、分析から除外した。

また、以下の分析では分布の差について言及し、検定を行うが²、本研究で扱う CSJ のデータは特に分節音の分析においてサンプル・サイズが極めて大きいため、検定力が過剰に高くなる。水本篤・竹内理(2010)は「p 値はサンプル・サイズ(標本数、サンプル数)が大きくなればなるほど、実質的な差がなかった場合でも、p 値は小さくなり、統計的に有意であるという結果が得られやすくなる」と指摘し、この問題への対処として、「サンプル・サイズによって変化しない、標準化された指標である効果量(effect size)」を示すことを勧めている。本研究では検定を行った際には検定量(3群以上の検定では多重比較(Mann-Whitney 検定)の検定量)から z 値を算出し効果量 r を算出して³示す。なお水本・竹内(2010)に、Mann-Whitney 検定において効果量は 0.1 で小、0.3 で中、0.5 で大という値が示されており、これに照らして効果量の大小を考える。

4.1. 鼻子音の持続時間の総合的分布

まず、CSJ に含まれる全音節頭鼻子音の持続時間についてその分布を示す。図1は鼻子音の持続時間を相対度数分布として示したものである。横軸は鼻子音の持続時間を 10ms ごとに区切った階級を示し、縦軸は各階級に属する鼻子音の出現頻度を表す。また表4に各鼻子音持続時間の分布について基本統計量を示す。

図1の分布および基本統計量から見て、鼻子音の持続時間の分布は、標準偏差±1の範囲を一つの手がかりとすれば、大体平均持続時間 51ms を中心として、おおよそ 30~70ms の持続時間のものが大半を占めるということが出来る。ただし持続時間の短くなることにはおのずと限界がある(0以下はありえない)のに対し、長くなる方にはそのような制約はないことから、わずかに持続時間の長い階級に裾の長い分布となっている。

今回の分析で得られた平均持続時間は先

データ数	127576
平均 (ms)	51.0
標準偏差 (ms)	18.7
最小値 (ms)	5
最大値 (ms)	549
尖度	13.96
歪度	1.31

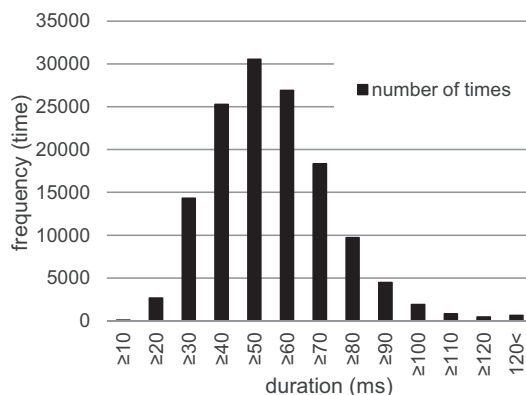


図1 CSJ 中の鼻子音 /m, n/ 持続時間の相対度数分布

行研究で示された値と比較すると、ポーズを含む長文読み上げによる Hiki et al. (1967) や 匂坂・東倉 (1984) の短文読み上げの結果に近い。先行研究の値は全体的に (Campbell (1992) を除いて)、比較的大きい値を示していると言える。特に Han (1962) や 匂坂・東倉 (1984) の /VCV/ 音韻連鎖の読み上げの結果は本校の結果からみてかなり大きい値を示している。本資料の結果および先行研究を含め、基本的に鼻子音の持続時間に関しては発話者の発話への注意の影響が反映され、注意が向くスタイルでの発話で、持続時間は長くなりやすいと考えることができる。

4.2. 発話速度による分布

先行研究によれば、発話速度の影響は子音にも観察されるものの、影響は子音よりも母音で大きいということであった (舛田・他 2002)。

次の図 2 は、発話速度と鼻子音の持続時間の関係を散布図で表したものである。発話速度 (モーラ/秒) はアクセント句 (以下、AP) を単位として計算した (モーラ数/AP 持続時間)。図 2 の横軸は当該鼻子音の含まれる AP の発話速度、縦軸は当該鼻子音の持続時間、一つのマーカーが一つの鼻子音を示す (次節の語の種類に関する分析との関連で、マーカーは語の種類ごとに分けて示している)。

まず AP の速度に関しては 30 モーラ/秒以下、そして鼻子音の持続時間は 250ms 以下におおよそおさまり、これら以上の値をとるものは例外的と言える。

先行研究で、発話速度の影響は、母音に対する方が大きいものの、子音長についても否定はされていない。今回の結果は鼻子音の持続時間に対して、発話速度が一定程度影響することを裏付けるものと見ることができる。すなわち全体に、発話速度が速くなるほど鼻子音の持続時間としては短いものが現れる。ただしこうした反比例の関係は持続時間の上限にだけあてはまるもので、各発話速度においてとる持続時間は、その上限以下の値で多様である。つまり発話速度の影響は、産出音声の持続時間の許容範囲の上限の変化として捉えることができる。発話速度が遅いほど鼻子音の持続時間の長いものが出現し得るが、逆に速度が速くなるほど短いものしか許されない。またその上限値は直線ではなく 2 次関数的なカーブを描いている。

なお長い (200ms 前後の) 持続時間の音声がよく見いだされるのは、発話速度 0 mora/sec 付近ではなく 5 mora/sec であり、短い (10ms 前後の) 持続時間の音声がよく見いだされるのは発話速度 30 mora/sec 付近ではなく 10 mora/sec である。つまり 5~15 mora/sec の速度の発話において最も多様な持続時間の音声が見られるといえる。

なお図 2 では鼻子音の含まれる語 (短単位) の種類 (自立語 (independent)、付属語 (dependent)、感動詞 (interjection)) について分けて示している。この語の種類と持続時間の

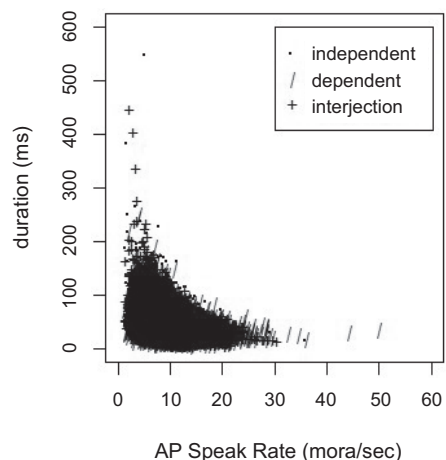


図 2 発話速度による持続時間の散布図

関係については次節で詳述するが、速度および鼻子音の持続時間に関係することをいえば、自立語と感動詞では比較的 AP 速度が遅く持続時間の長いものが観察され、一方付属語は AP 速度の速く持続時間が短いものが比較的多く観察されるようである。

4.3. 語（短単位）の種類による違い

本節では鼻子音の含まれる語（短単位）の種類による鼻子音の持続時間について述べる。ここで区別する語の種類とは自立語、付属語、それと感動詞の3つである。Kaiki & Sagisaka (1992) では、自立語と付属語の間で子音の持続時間に違いが見られ、全般に自立語では伸長が起き付属語では短縮が起きると指摘している⁴。ここでは CSJ の品詞情報のラベルで「名詞、動詞、形容詞、形状詞、副詞、代名詞、連体詞、記号」を自立語とし（図3の independent）、また「助詞、助動詞、接続詞、接頭辞、接尾辞」を付属語とする（図3の dependent）して、各種に含まれる鼻音の持続時間を比べる。また、感動詞はその具体的な発話内容としてはいわゆるフィラー（CSJ で (F) の記号が付されている）がほとんどを占め（99.9%）、ここでは独立した1項目とする（図3の interjection）。

図3は各項目における鼻子音の持続時間の相対度数分布を示したものである。横軸は持続時間の階級、縦軸は各階級のデータの割合を示す。また表5にデータ数と平均、標準偏差を示す。

図3を見ると、付属語<自立語<感動詞の順に持続時間の長い鼻音子音が現れやすい傾向を見ることができる。この傾向（付属語<自立語）については先行研究の指摘を支持するものであり、この点は一貫している情報として捉え得る。ただし、語の種類が持続時間に影響を与えてとしても、その差は約3ms ずつ、付属語と感動詞の間でも約6ms の違いで、非常に小さな違いである。

これらの語の種類の違いによる持続時間に関して Kruskal-Wallis 検定を行ったところ 0.1%水準で有意差が確認された ($\chi^2=1065.03$, $df=2$, $p\text{-value} < 2.2e-16$)。また多重比較を行ったところ、0.1%水準ですべての項目間で有意差が確認された (Bonferroni, 全ての項目間で $p < 2e-16$)。ただし多重比較 (Mann-Whitney 検定) の検定量から効果量 r を求めたところ、すべての項目間で $r < 0.1$ であった。従って語の種類間に見られた差はあまり重視すべきではないと考えられる。

表5 語の種類別平均・標準偏差

	自立語	付属語	感動詞
データ数	51660	63801	12115
平均 (ms)	52.0	49.4	56.4
標準偏差 (ms)	18.9	17.5	25.3

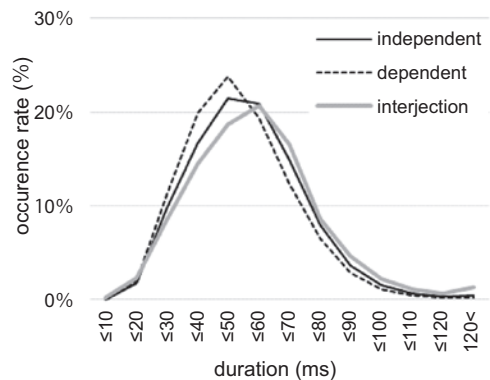


図3 語の種類別相対度数分布

4.4. 調音位置による分布

先にも述べたように、先行研究によれば子音の持続時間を決める大きな要因はその音素の種類であるという。日本語の音節頭子音の音素としては両唇音 /m/ と歯茎音 /n/ がある。

CSJ の分節音ラベリングでは、鼻子音のラベリングとして /m/ に対して「m, my (マ行拗音の子音)」、/n/ に対して「n, ny (ナ行拗音の子音)、nj (母音イの前で口蓋化した n)」が付されている。本稿ではこのうち「m」と「my」は主たる調音位置は変わらないものとして捉え、どちらも [m] として分析する(後続母音による影響は 4.5.3 節で述べる)。一方、「nj」と「ny」は [n] として、「n」は [n] として主たる調音位置が異なるものとして分けて扱う。従って本節の分析では [m, n, ɲ] を調音点を比較する分析項目とする。なお /g/ の異音として [ɲ] が産出される可能性もあるが、CSJ で现阶段ではラベリングされておらず今回は分析対象としない。

図 4 は各調音位置の鼻子音の持続時間を相対度数分布として示したものである。横軸は鼻子音の持続時間を 10ms ごとに区切った階級、縦軸は各階級に属する鼻子音の出現割合を表す。また表 6 に平均と標準偏差を示す。

結果を見ると、違いは /m/ と /n/ の音素間に見られ、同じ音素に属す異音 [n] と [ɲ] の間にはほとんど見られない。調音位置の違いに関して Kruskal-Wallis 検定を行ったところ 0.1% 水準で有意差が確認され ($\chi^2=12926.41$, $df=2$, $p\text{-value} < 2.2e-16$)、またその結果について多重比較 (Bonferroni) を行ったところ、0.1% 水準で [m] とそれ以外の調音位置との間には有意差が確認されたが、[n] と [ɲ] の間には確認されなかった (Bonferroni, $p([m]:[n], [m]:[ɲ]) < 2e-16$, $p([n]:[ɲ])=0.65$)。また多重比較の検定量から効果量を算出すると、[m] と [n] 間で中程度 ($r=0.32$)、[m] と [ɲ] の間で小程度 ($r=0.28$) の効果量が確認されたが、[n] と [ɲ] の間では $r < 0.1$ であった。以上のことから、やはり [m] とそれ以外の調音位置の間のみ、持続時間の差を認めるべきと結論できる。すなわち、両唇音では他の調音位置に比べ、持続時間が長くなる傾向があると言える。

[m] と、[n] あるいは [ɲ] との調音位置の違い (両唇と口蓋) は、[n] と [ɲ] の間の違い (歯茎と硬口蓋) より大きく、この違いが反映されている可能性がある。例えば口腔内の呼気圧が一定程度に高まり閉鎖が開放するまでの所要時間の影響、あるいは調音器官の弾性や動作機構の違いの影響などを考えることができよう。

表 6 調音位置別平均・標準偏差

	[m]	[n]	[ɲ]
データ数	51615	63683	20823
平均 (ms)	57.7	46.5	47.2
標準偏差 (ms)	19.1	17.3	18.8

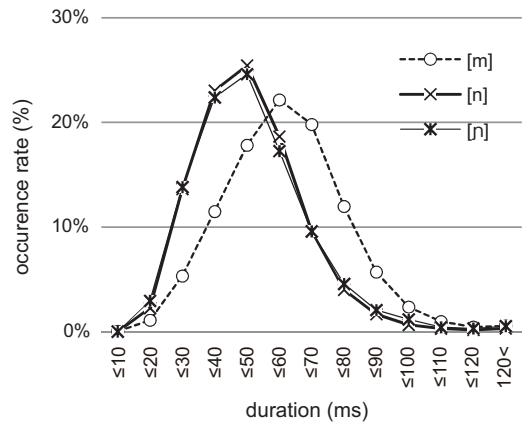


図 4 調音位置別相対度数分布

4.5. 隣接音による分布

鼻子音の隣接音に関する音環境については、先行音と後続音に分けて考察する。まず当該鼻子音の先行音の有無 (=ポーズの有無)、また先行音がある場合その種類、そして後続音の種類による持続時間への影響を見ていく。

4.5.1. 直前のポーズ

本節では鼻子音の直前のポーズの有無による鼻子音の持続時間への影響を分析する。CSJでは各分析単位に開始時間と終了時間がラベリングされている。従って先行分節音の終了時間と当該鼻子音の開始時間の差が0であれば「ポーズ無 (-pause)」、マイナス値であれば「ポーズ有 (+pause)」とすることができる。なおCSJの補助ラベリングで呼吸音を表す はポーズ有として扱った。

図5は各条件下の持続時間の相対度数分布を示したものである(各軸の内容はこれまでと同様)。*/m/*と*/n/*の間に興味深い関係が見られるためこれらを分けて示す。表7にデータ数、平均、標準偏差を示す。

前節で*/m/*と*/n/*の間に持続時間の違いが見られることを指摘したが、図5から、これが、ポーズのない場合に限られることがわかる。

直前にポーズがある場合には音素間の違いはほとんどなく、持続時間の分布のピークは比較的持続時間の短い階級にあるが低く、分布が全体にばらついて多様な持続時間の音声が見られることがわかる。これに対し、直前にポーズがない場合には、相対的にばらつきが小さく、比較的持続時間の長い階級に分布のピークがある。そして調音位置による違いが大きい。

このポーズと調音点の組み合わせによる4項目の分布の差についてKruskal-Wallis検定を行ったところ⁵、0.1%水準で有意差が確認された($\chi^2=15424$, $df=3$, $p\text{-value} < 2.2e-16$)。なおその結果について多重比較(Bonferroni)を行ったところ全ての項目間で0.1%水準で有意差が確認されたが、効果量については*/n/* -pauseと*/m/* -pauseとの間で中程度($r=0.35$)、また*/n/* +pauseと*/m/* -pauseの間($r=0.13$)、および*/m/* +pauseと*/m/* -pauseの間($r=0.19$)で小程度認められたが、その他の項目間では0.1未満であった。

以上の結果から、まず、ポーズの有無は鼻子音の持続時間に影響を与え得ると言える。ただしその内容は鼻子音の調音位置によって異なり、*/n/*では持続時間の長いものが現れにくくなるのに対し、*/m/*では現れやすくなると言える。言い換えれば、調音位置による持続時間の分布の違いは話中の場合に限られるとも言える。

表7 直前のポーズの有無別平均・標準偏差

鼻子音 ポーズ	<i>/n/</i>		<i>/m/</i>	
	あり	なし	あり	なし
データ数	4244	80260	6942	44667
平均(ms)	51.9	46.4	49.7	59.0
標準偏差(ms)	25.4	17.2	23.9	17.9

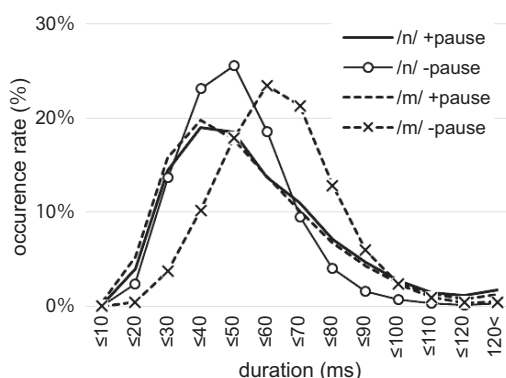


図5 直前ポーズ有無別相対度数分布

4.5.2. 先行音

先行研究によれば、子音は母音に比べて隣接音の影響を受けにくい(2.1節参照)。本節では特に先行音について、自然発話における持続時間への影響を確かめる。鼻子音の先行音としては、本稿の資料では母音と撥音 /N/ の他、促音 /Q/ と CSJ の補助ラベル⁶ を用いてコーディングされた音声を観察された。ただし、先述のとおり /N/ は分析から除外し、また /Q/ については 23 例抽出されたがそのすべてが言いよどみなど語彙的なものではないことを確認し(例「このもっ物語の」)、また補助ラベルによるものについては分析上の明確性を確保するため、分析対象から除外した。結果としてここで分析対象とする音節頭の鼻子音の先行音は母音のみである。

日本語は /i, e, a, o, u/ の母音音素を持ち、それぞれにおいて長短の音韻的区別がある。ここではまず 5 母音間の比較を行い(図 6)、次に長短による比較を行う(図 7)。図はそれぞれ鼻子音の持続時間について、相対度数分布で表したものである(図の見方はこれまでと同じ)。表 8 に平均と標準偏差を示す(表中 V は母音、: は長音を表す)。

図 6 を見ると、母音間でわずかながら分布に違いがあるように見える。先行する 5 母音間の違いについて Kruskal-Wallis 検定を行ったところ 0.1% 水準で有意差が確認された ($\chi^2=1262.48$, $df=4$, $p < 2.2e-16$)。また多重比較(Mann-Whitney 検定, Bonferroni)を行ったところ /a/ と /u/ の組み合わせを除くすべての組み合わせで 0.1% 水準で有意差が確認された ($p (/a/:/u/) = 0.087$, $p (/i/:/e/) < 3.3e-06$, $p (/o/:/u/) = 2.9e-13$, その他の組み合わせで $p < 2e-16$)。ただし多重比較の結果に対して効果量を算出したところ、/i/ と /o/ ($r=0.11$)、/u/ と /e/ ($r=0.11$)、/e/ と /o/ ($r=0.13$) で小程度の効果量が認められたが、その他の項目間では r は 0.1 未満だった。

有意差と小程度の効果量の確認された /i/ と /o/、/u/ と /e/、/e/ と /o/ の間には調音における舌の

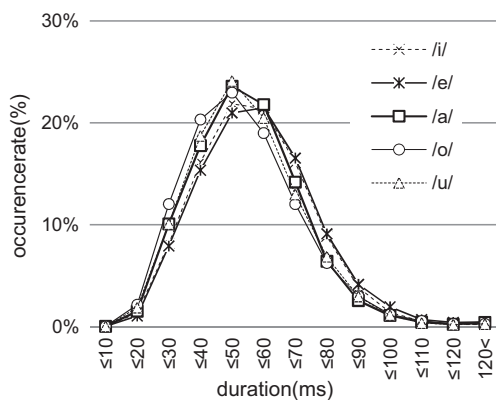


図 6 先行音 (5 母音) 別相対度数分布

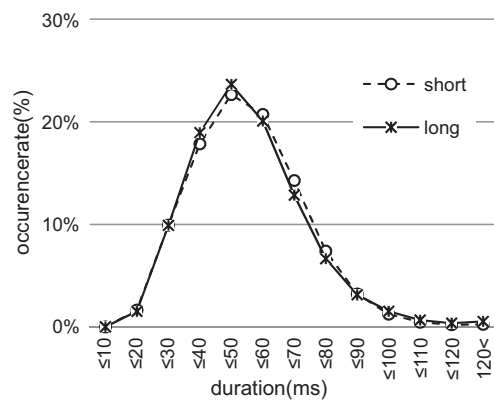


図 7 先行音 (母音長短) 別相対度数分布

表 8 先行音別平均・標準偏差

先行音	/i/	/e/	/a/	/o/	/u/	V	V:
データ数	28147	13454	24416	34189	17151	102678	14679
平均 (ms)	52.8	54.1	50.6	49.2	50.1	51.0	51.1
標準偏差 (ms)	17.6	19.2	18.0	18.1	17.4	17.9	19.2

盛り上がりの前後との相関を見ることができる。すなわち後舌母音 /o, u/ よりも前舌母音 /i, e/ が先行する場合の方が鼻子音の持続時間の長いものが出現しやすい傾向にあるといえる。しかし一方で /i/ と /u/ の間には見られないため、これをどう解釈すべきか難しい。また、その平均値の違いを見ると、その差は3~5ms と小さいものであり、もしこれらの母音間の違いが一定しているとしても、それ程重要な違いとして、取り上げるべきものではないと考える。

次に先行母音の長短の影響について見てみよう。図7を見ると長母音と短母音の分布にほとんど差がない。この分布の違いについて Mann-Whitney 検定を行った結果、有意水準 0.1% で有意差が確認されたが (W = 743160000, p = 0.007)、効果量は極めて小さいので (r = 0.008)、先行母音の長短は鼻子音の持続時間に影響を及ぼすとは考えるべきでないだろう。

以上のことから鼻子音の持続時間に関して、先行音の影響は、その有無が重要であり、その音種の影響はあまり大きくないと考えられる。

4.5.3. 後続音・音節の軽重

本節では隣接音のうち後続音に関して検証する。また鼻子音に後続する母音の長短は同時にその鼻子音を含む音節の音韻的な軽重の問題とも関係するため、これを含めて分析する。ただしここで扱うのは、当該鼻子音に対し直接後続する母音とそれに続く音、すなわち後続の2モーラまでである。

音節頭鼻子音に直接後続するのは母音である。まず5母音の比較を行った後(図8)、母音の長短をふくむ音節の軽重による鼻子音の持続時間について見る(図9)。図はそれぞれ鼻子音の持続時間について相対度数分布で表したものである(図の見方はこれまでと同じ)。表9に平均

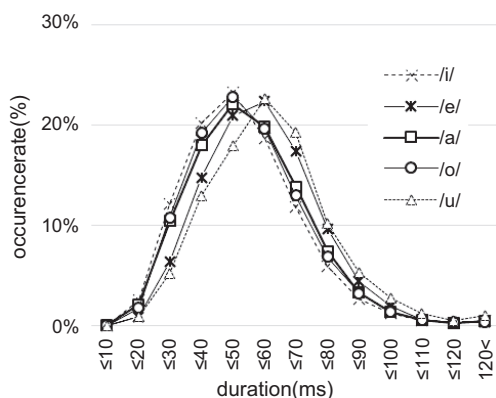


図8 後続母音別相対度数分布

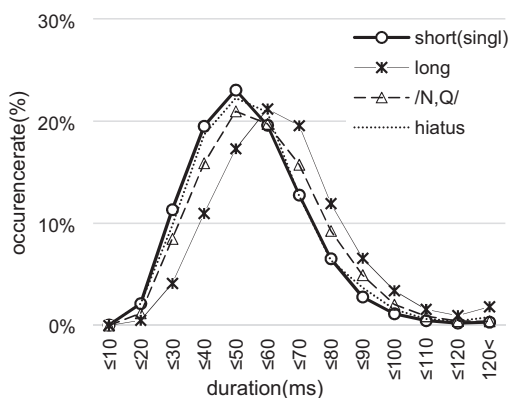


図9 音節の軽重別相対度数分布

表9 後続母音別平均・標準偏差

後続音	/i/	/e/	/a/	/o/	/u/
データ数	25446	8583	46324	52487	3281
平均 (ms)	51.1	54.7	51.0	50.5	57.7
標準偏差 (ms)	19.0	18.1	19.0	19.0	19.7

表10 音節の軽重別平均・標準偏差

音節種	短母音	長母音	/N, Q/ 音節	母音連続
データ数	108217	9046	6568	12290
平均 (ms)	49.6	61.0	51.5	54.0
標準偏差 (ms)	18.2	23.3	21.3	19.1

と標準偏差を示す。図9中 short/long は母音の短/長、/N, Q/ は後続母音後に撥音また促音が接続する音節、hiatus は後続母音後に AP 区切りをはさまず母音が連続するもので、ここでは母音連続と呼ぶ。ただしここでの母音連続のデータは 1 形態素内に実現するものに限られていない。

図8を見ると、/i, a, o/ と /e, u/ の間で違いがあるように見え、最頻値で 1 階級 (10ms) の違いがある。後続母音に関して Kruskal-Wallis 検定を行ったところ 0.1% 水準で有意差が確認され ($\chi^2=1280.6$, $df=4$, $p < 2.2e-16$)、多重比較 (Mann-Whitney 検定、Bonferroni) でもすべての項目間で有意水準 0.1% で有意差が確認された (p (/a/ : /o/) = $8.0e-07$, p (/u/ : /e/) = $9.4e-12$, その他の組み合わせで $p < 2e-16$)。ただしこの結果に対し効果量を見てみると、/i/ と /u/ ($r=0.15$)、/i/ と /e/ ($r=0.15$) で小程度認められたが、その他の項目間では r は 0.1 未満だった。

/i/ と /e, u/ との間の違いは統計的には一程度認められるものではあるが、しかし後続母音の開口度や舌の前後位置といった調音特徴から一貫した解釈することはできない。効果量から考えて、この有意差はそれほど重視すべきではないと思われるが、他の要因が隠れている可能性もあり、今後の検討を要する。

次に音節の軽重と鼻子音持続時間との関係を見てみよう (図9)。ここでは後続母音が、短母音 = 母音連続 < /N, Q/ 音節 < 長母音音節の順で、鼻子音の持続時間が長くなる傾向が見られ、短母音と長母音の間では平均値で 10ms 程度の違いが見られる。これらの違いについて Kruskal-Wallis 検定を行ったところ 0.1% 水準で有意差が確認され ($\chi^2=2999.7$, $df=3$, $p\text{-value} < 2.2e-16$)、多重比較 (Mann-Whitney 検定、Bonferroni) でもすべての項目間で有意水準 0.1% で有意差が確認された (p (短母音 : /N,Q/ 音節) = $9.7e-09$, それ以外の組み合わせでは $p < 2e-16$)。ただし効果量を算出したところ長母音音節と他の音節の間には小程度の効果量が認められたが (p (長母音 : 短母音) = 0.15, p (長母音 : /N,Q/) = 0.24, p (長母音 : 母音連続) = 0.16)、その他の項目間では 0.1 未満だった。

この結果からは、音節の軽重が鼻子音の持続時間にある程度関わると見える。ただしこれも効果量からみたとおり、それほど大きくとりあげるべき要因とは言えない。以上のことから、隣接音に関しては、鼻子音の持続時間に与える影響は部分的であることがうかがえる。この結果は子音は母音ほど隣接音の影響は受けないという先行研究の指摘 (2.1 節) と一致するものである。

4.6. AP 内の位置による違い

先行研究によれば、AP 頭では子音の持続時間が伸長するが、ただし鼻子音は比較的その影響は受けにくいという (Kaiki & Sagisaka 1992)。ここでは AP の第 1 モーラとその他モーラ (第 2 モーラ以降) とで鼻子音の持続時間を比較する。なお AP 第 1 モーラは、その前にポーズが入るケースも多いと考えられる。ポーズの先行は鼻子音の持続時間の伸長に影響することがすでに明らかのため (4.5.1 節) 分けて考える。

図9はそれぞれ鼻子音の持続時間について、相対度数分布で表したものである (図の各軸の内容はこれまでと同じ)。第 1 モーラに位置する場合、やはりポーズが先行するか否かによって持続時間の分布は異なり、AP 内の位置の影響を見るにはポーズが先行しないもの同士を比べる必要があることがわかる。全体に、ポーズが先行する場合には持続時間の分布のピークは小さい階

級（持続時間が短い）にあるが低く、多様な持続時間のもが現れやすいことは既に述べた通りである（4.5.1節）。またポーズの先行しない第1モーラの鼻子音（1st mora -pause）と第2モーラ以降のモーラ（other mora）の分布を比べると、第1モーラである場合に持続時間の長いものが現れやすいことがわかる。この結果は上記の先行研究の指摘と一致する。

このポーズと AP 内位置の組み合わせによる3項目の分布について Kruskal-Wallis 検定を行ったところ⁵、0.1%水準で有意差が確認された ($\chi^2=4504.2$, $df=2$, $p\text{-value} < 2.2e-16$)。なおその結果について多重比較 (Bonferroni) を行ったところ全ての項目間で 0.1%水準で有意差が確認されたが (p (1st mora pause-: other mora)= $4.7e-16$ 、その他の組み合わせで $p < 2e-16$)、効果量については

1st mora pause- とその他の項目との間で小程度認められたが (r (1st mora -pause: 1st mora +pause)=0.25, r (1st mora -pause: other mora)=0.19)、その他の項目間では 0.1 未満であった。

検定の結果からは、この AP 内位置（およびポーズ）はそれ程重要な要因として取り上げられるものとは言えない。しかし、この AP 内位置の持続時間の影響については、句頭音調の上昇にかかる時間が関わっている可能性が考えられる。ポーズが先行する場合には、発音の初頭からある程度高く始めることが可能であるが、ポーズがない場合には先行 AP 末の高さからの上昇が必要となる。このため、上昇にかかる時間が長くなり、鼻子音の持続時間もそれに伴って長くなる場合があると考えられる。

5. 結論

以上の結果を整理する。まず CSJ における鼻子音の持続時間の分布は平均値が 51ms で、おおよそ 30~70ms の持続時間のもが大半を占めていた。また持続時間の分布への影響を与える要因として、AP 発話速度、調音位置、直前のポーズの有無、音節の軽重、AP 内の位置（第1モーラか否か）が関わっている子が見られた。これらの要因と持続時間の関係については、おおむね先行研究の指摘と一致していたと捉えることができる。

一方で、分析結果をヒストグラムの形で表現したことにより見えた分布上の特徴もあったように思われる。例えば直前のポーズの有無については、ヒストグラムとしてデータを表現した際に、分布の平均値やピークの大小だけでなく、ばらつきや歪みなどの違いも目立った。ポーズが先行する場合、その分布のピークは相対的に短い持続時間にあるが、しかし一方で分布の裾が長く持続時間の長いものも他の条件に比べ出現割合が高い。ポーズがないという条件に比べ様々な

表 8 AP 内位置別平均・標準偏差

	第1モーラ		その他モーラ
	ポーズ有	ポーズ無	
データ数	10024	18114	107983
平均 (ms)	49.9	59.9	49.4
標準偏差 (ms)	24.4	21.7	17.5

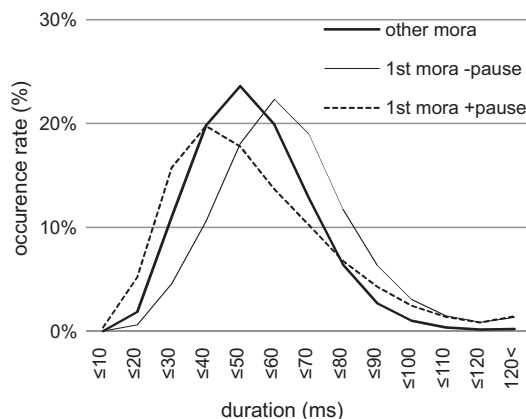


図 9 AP 内位置別相対度数分布

持続時間の音声許容されることとして解釈ができる。

以上の結果を総合すると、鼻子音の持続時間について次のように考えられる。まず発話速度により、発話全体の流れの中での当該子音に与えられる時間が決定される。調音に必要なとする時間は基本的には調音位置で異なるが、ポーズが先行する場合には、調音位置の違いに関係なく、多様な持続時間で実現される。またポーズをはさまずに先行 AP に当該 AP が続く場合、その第 1 モーラ (AP の頭子音) において、鼻子音はおそらく AP 頭のピッチ上昇と関わって、与えられた速度の枠の中で伸長する場合がある。また長母音が後続する場合はわずかに子音が伸長する傾向がある。

本稿では自発音声という点で先行研究で扱われた資料とは異なる性格の音声資料を分析してきた。自発音声コーパスにおける鼻子音の持続時間は先行研究で示された結果の中でも、発話者の意識が向きにくい資料を用いた結果と類似していた。筆者は、本稿の最初に述べたように、日本語の鼻子音の持続時間について、地域的・世代的多様性を視野に入れた資料を収集・分析するつもりである。その際、自然発話でない資料を用いて測定した場合には、自然発話の場合よりも長いものが出現すると推測することができる。また東京を中心とする話者の結果を基準に対比させることも可能である。

注

- ¹ 鼻音の前に直接接する音とする。
- ² R version 3.0.2 を使用した。また検定はすべてノンパラメトリックな手法を採用した。
- ³ $Z = (|U - E(U)|) / \sqrt{V(U)}$ 、 $E(U) = n_1 n_2 / 2$ 、 $V(U) = (n_1 n_2 (n_1 + n_2 + 1)) / 12$ 、 $r = Z / \sqrt{n_1 + n_2}$ 。ただし U は多重比較で算出された検定量。
- ⁴ なお Kaiki & Sagisaka (1992) では自立語の中で数量詞や固有名詞での伸長も指摘している。CSJ では固有名詞と数詞を短単位の付加情報の内「その他」にラベリングしているが、今回はこの情報を利用することができなかった。今後の課題である。
- ⁵ 本来は交互作用を含めた検定を行うべきであるが、等分散性が確認できず、二元配置の分散分析に対応するノンパラメトリック法はないためこの検定法を採用した。
- ⁶ 先行音として検出された補助ラベリングは次のものである。<uv>、<sv>、<fr>、<fv>、<?>。これらの具体的内容については CSJ のマニュアルを参照のこと。

【謝辞】 本稿執筆の際には多くのコメントをいただきました。また、本稿で使用したデータを CSJ から抽出する際には前川喜久雄先生にご協力いただきました。記して改めて感謝致します。

参考文献

- 板橋秀一 (2005) 『音声工学』 森北出版
- 国立国語研究所 (2006) 『日本語話し言葉コーパスの構築法』 (国立国語研究所報告 124) 国立国語研究所
- 匂坂芳典・東倉洋一 (1984) 「規則による音声合成のための音韻時間長制御」『電子通信学会論文誌』 J67-A (7), 629-636.
- 高田三枝子 (2011) 『日本語の語頭閉鎖音の研究—VOT の共時的分布と通時的変化—』 くろしお出版
- 前川喜久雄 (2004) 「『日本語話し言葉コーパス』の概要」『日本語科学』 15, 111-133.
- (2010) 「日本語有声破裂音における閉鎖調音の弱化」『音声研究』 14(2), 1-15.

- (2011) 「『日本語話し言葉コーパス』を用いた自発音声の分析」『情報処理学会研究報告』2011-CH-92(4), 1-4.
- 舩田剛志・戸田智基・川波弘道・猿渡洋・鹿野清宏 (2002) 「発話速度の異なるデータベースを用いた音声合成手法の検討」『電子情報通信学会技術研究報告』(音声) 101(603), 61-62.
- Beckman, Mary (1982) "Segment duration and the 'mora' in Japanese." *Phonetica* 39, 113-135.
- Campbell, Nick (1992) "Segmental elasticity and timing in Japanese speech." Yoh'ichi Tohkura, Eric Vatikiotis-Bateson and Yoshinori Sagisaka (eds.) *Speech perception, production and linguistic structure*, Ohmsha, IOS Press, 403-418.
- Han, Mieko (1962) "The feature of duration in Japanese." 『音声の研究』 10, 65-80.
- Hiki, Shizuo, Yoshinari Kanamori, and Juro Oizumi (1967) "On the duration of phonemic segment in sentence speech." 『日本音響学会誌』 23(5), 314-317.
- Kaiki, Nobuyoshi and Yoshinori Sagisaka (1992) "The control of segmental duration in speech synthesis using statistical methods." Yoh'ichi Tohkura, Eric Vatikiotis-Bateson and Yoshinori Sagisaka (eds.) *Speech perception, production and linguistic structure*, Ohmsha, IOS Press, 391-402.
- Keating, Patricia (1984) "Phonetic and phonological representation of stop consonant voicing." *Language*, 60(2), 286-319.