

## 鼻音調音時のエネルギー遷移曲線における話者の個人性\*

◎網野加苗, 荒井隆行 (上智大・理工)

## 1 はじめに

個人性の知覚について調べることは、法科学などにとって有益である。話者認識聴取実験の結果は個人性に関わる音響特徴量を知る手がかりにもなるため、音声工学にも貢献できる [1]。

聴取による話者識別においては、呈示する刺激音の発話内容によって正答率が影響されることが知られているが [2], 著者らの今までの研究において、鼻音を含む音節が口音のみの音節よりも高い正答率を得ることが一貫して示されてきた [3-7]。

前回の報告においては後続母音の影響を調べ、後舌母音の有効性が示されたが [7], 子音に関してはやはり鼻音が有効であることが明らかになった。また同実験の刺激音について、音節内のエネルギー遷移の曲線が話者によって異なる形状を示し、話者間の差が鼻音のみで有意であることも分かった [8]。

本稿では、さらに話者数を追加してエネルギー曲線の分析を行ったので報告する。

## 2 分析

## 2.1 パラメータ

前研究 [7, 8] に倣い、エネルギーの時間変化を分析パラメータとした。このパラメータは、調音運動を反映しており、話者の発話動作における癖が表れる可能性がある。

分析条件をTable 1に示す。フレーム長 30 ms, シフト長 10 ms でエネルギーを計算した。

## 2.2 音声資料

前研究と同様、本研究においても電子協日本語共通音声データ [9] より、成人男性の単音節発話をを用いた。分析対象の音節頭子音は、/d/ /t/ /z/ /s/ /m/ /n/ の6種類とし、閉鎖音は3つ、摩擦音と鼻音はそれぞれ4つと5つの後続母音のサンプルを対象とした。

前研究にて分析を行った4名 (#1-4) に加え、今回新たに6名 (#5-10) の発話音声を採用し、計10名分のデータについて結果を考察

する。また、各音節につき各話者の3回分の発話を分析した。

## 2.3 傾きの計算

フレームごとのエネルギーの時間変化を曲線にし、その遷移曲線の左側のカーブの傾きを求めた。この区間は音節頭子音から核母音へのエネルギー上昇を示しており、頭子音の調音方法によって遷移の仕方が異なることが予想される。また、子音別の話者性もこの区間に現れる可能性が高い。

ここでは、エネルギーの値を音節内の最大値で正規化し、0.1~1.0の範囲の傾きを求めた。

## 3 結果

10名の話者の鼻音を含む音節 /mV/ /nV/ の遷移曲線をFigure 1に示す。数名の話者ではばらつきがあるものの、概ね話者ごとに異なる形状を示しており、また話者内での再現性が高い。

曲線の傾きについて、調音方法別に話者ごとの平均値とF比の値をTable 2に示す。F比は、話者内分散に対する話者間分散の比で、値が大きいほど個人性を多く含む可能性を示している。Table 2から、鼻音のエネルギー曲線の傾きのF比が口音と比べて大きいことが分かる。ここで、鼻音と口音（閉鎖音および摩擦音）の差はマン・ホイットニーのU検定において有意であった ( $p < .05$ )。

## 4 考察

CV単音節の発話において、エネルギー遷移曲線は話者の子音調音時の一特性を反映している。前研究の4名から話者を10名に増やした場合でも、特に鼻音において、曲線の傾きの個人性が大きいことが分かった。

Table 1 Parameters for the energy analysis

Sampling frequency	16 kHz
Resolution	16 bit
Frame length	30 ms
Frame shift	10 ms

\* Speaker individualities in the energy contours of nasal articulation, by AMINO, Kanae and ARAI, Takayuki (Sophia University).

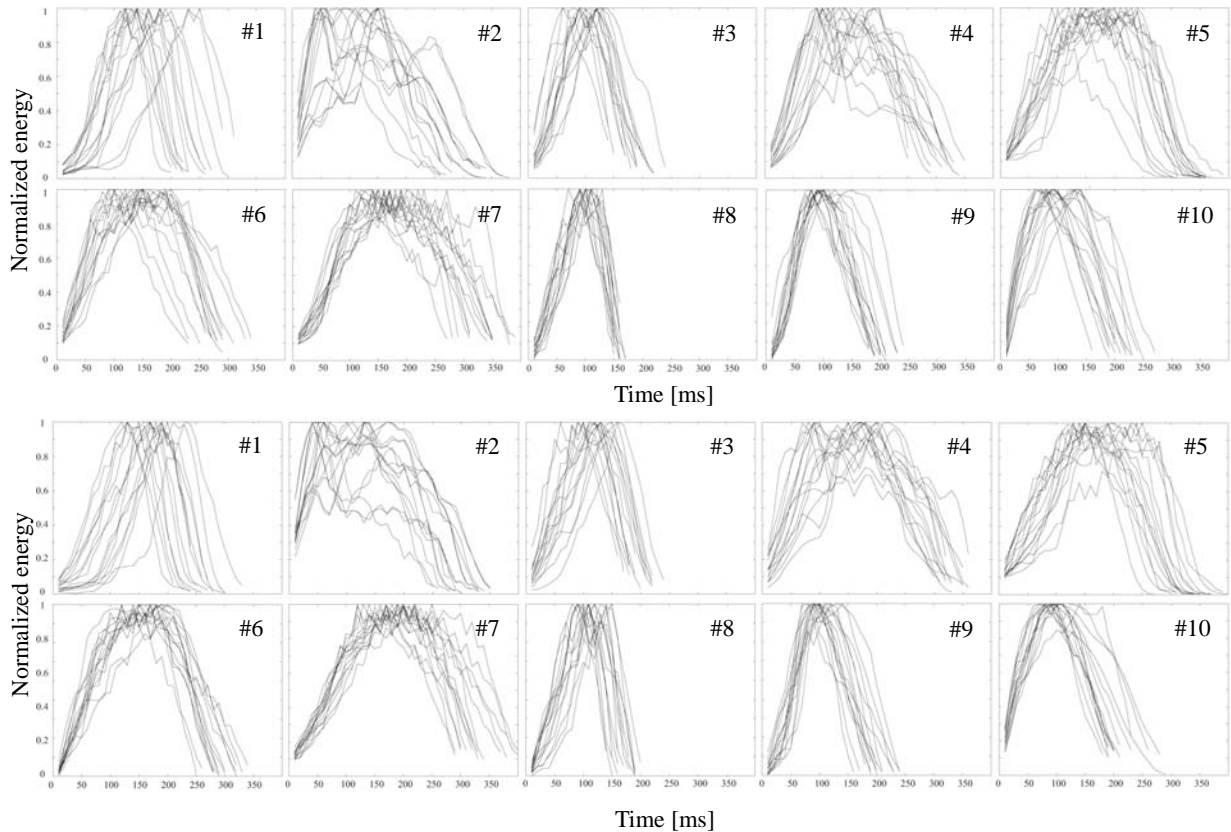


Figure 1 Energy contours; /mV/ (upper panels) and /nV/ (lower panels) syllables uttered by ten speakers

Table 2 Mean slope values and *F*-ratios for ten speakers' utterances of stops, fricatives and nasals

Speakers		#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Ave.
mean slope [dB/ms]	stops	0.128	0.107	0.132	0.145	0.082	0.081	0.081	0.115	0.148	0.158	0.118
	fricatives	0.161	0.114	0.129	0.130	0.112	0.088	0.008	0.156	0.177	0.149	0.130
	nasals	0.086	0.105	0.103	0.075	0.059	0.063	0.057	0.095	0.109	0.094	0.085
<i>F</i> -ratio	stops	0.597	0.683	0.929	0.993	1.601	2.16	1.361	5.431	1.669	1.172	1.660
	fricatives	0.813	1.784	1.659	0.800	1.051	2.958	0.988	2.056	0.881	1.204	1.420
	nasals	2.829	0.313	1.022	1.437	3.398	3.376	5.437	3.808	2.266	1.876	2.612

本稿で用いたCV遷移区間のエネルギー曲線の傾きは、話者性を示す音響特徴量として単独で用いられるパラメータではないかもしれないが、話者を分類する手がかりとしては使えると考えられる。Figure 1の鼻音の曲線を見ても、10名の話者をいくつかのパターンに分類することができる。本稿ではエネルギーが 0.1 から 1.0 までの区間について傾きを計算したが、今後スペクトル分析と併せて子音部など特定の区間について計算すれば、より明確な話者間差が出るかもしれない。

鼻音で特に曲線の傾きの話者間差が大きかった点については、鼻音が軟口蓋制御を伴うことが挙げられる。Figure 1から曲線がピークに達するまでの時間を読み取ると、100 ms 以内であることが分かる。軟口蓋挙上に関わる筋肉は随意筋であるが、この短時間で話者が

意図的な制御を行える可能性は低い [10]。

また、個人性知覚との関連については、聞き手がエネルギー遷移における話者性を手がかりとしているかどうかは不明であるため、今後実験によって検証する必要がある。

#### 謝辞

この研究は上智大学オープン・リサーチ・センターの助成を得た。

#### 参考文献

- [1] O'Shaughnessy, "Speech Communications," Addison-Wesley Publishing Co., 2000.
- [2] Bricker and Pruzansky, in Lass [ed.], "Experimental Phonetics," 295-326, Academic Press, 1976.
- [3] 網野, 信学技報 SP, 104, 49-54, 2004.
- [4] Amino *et al.*, Acoust. Sci. Tech. 27, 233-235, 2006.
- [5] 網野ら, 信学技報 SP, 105, 109-114, 2006.
- [6] Amino and Arai, Acoust. Sci. Tech., 28, 128-130, 2007.
- [7] 網野, 荒井, 音講論(春), 443-444, 2008.
- [8] Amino and Arai, Proc. Interspeech (*to appear*), 2008.
- [9] JEIDA Japanese Common Speech Data Corpus [http://www.sunrisemusic.co.jp/dataBase/fl/voicebase01\\_fl.html](http://www.sunrisemusic.co.jp/dataBase/fl/voicebase01_fl.html)
- [10] Hollerbach, Trends in Neuroscience, 5, 189-192, 1982.