

## LPC 残差信号による子音認識と音素配列

小松雅彦 (上智大学理工学部・アルバータ大学大学院言語学科)

新谷敬人 (上智大学大学院外国語学研究科)

高澤美由紀 (清泉女子大学)

荒井隆行 (上智大学理工学部)

<http://www.splab.ee.sophia.ac.jp>

### 1. はじめに

LPC 残差信号による子音の認識を調べるために聴取実験を行った。その結果に対して音素配列的影響を検討し、さらに音韻的素性による分析を試みた。

Mori et al. (1999)では、日英語の言語識別における韻律情報の役割を調べるために、自発的発話から作成した LPC の残差信号を用いて聴取実験を行った。残差信号のインテンシティーは原音声と同じになるよう調整されていた。実験後のアンケートでは、用いられた刺激音からはスペクトル包絡情報が除去されていたにもかかわらず、単語が聞き取れたという回答が寄せられた。Mori et al.は、倍音構造やノイズ成分の有無、インテンシティーの時間的変化等という残差信号に含まれる不完全な形での音素配列的信息が、被験者に単語を聞き取らせた、あるいは聞き取ったと錯覚させた可能性を指摘している。

残差信号から、子音の識別が偶然以上の確率でできることは、CV 音節を刺激音とした聴取実験でわかっている (Arai et al., 1999; Komatsu et al., 2000)。そこで、筆者らは、音素配列的信息がどのように残差信号による子音認識に影響を及ぼすかを調べることにした。音素配列的な制約は言語によって異なるため、日本語のほか英語・スペイン語についても調べることにした。

LPC 分析は、音声信号を音源情報 (残差信号) とフィルタ情報 (LPC 係数) に分離する。これを単純に、超分節素性と分節素性の分離ととらえることはできない。本研究では、「音源情報」と音韻的素性との対応を検討するために、実験結果の音韻的素性による分析も試みた。

### 2. 実験

インテンシティー情報を保った LPC 残差信号を刺激音として、音節頭の子音連鎖中の子音の認識率を調べた。音素配列的制約との関係を調べるために、日本語・英語・スペイン語の 3 言語について実験を行った。

#### 2.1 資料

表 1 に示すような子音または子音連鎖を語頭に持つ語を用いた (子音・子音連鎖の種類 1 つにつき 1 語)。日本語では音節頭に子音連鎖が

来ないので、 $C_y$  および  $C_1V_0C_2$  を資料として加えた。 $C_1V_0C_2$  の  $V_0$  は音韻的無声化した母音で、 $C_1$  は無声閉鎖音、 $C_2$  は無声閉鎖音か無声摩擦音、 $V_0$  は狭母音である (Tsuchida, 1998)。(母音の脱落ととらえられるもの [前川, 1988; など] は、含まれていない。)

音節の種類や子音の分類については、日本語は猪塚・猪塚(1993)、英語は御園(1995)、スペイン語は出口(1997)、Canellada y Madsen (1987)、Quilis (1988)を参考にした。

表 1 資料

	語頭の子音連鎖	連鎖の種類 (語数)	語の長さ
日本語	C (C は直音)	14	2~3 モーラ
	$C_y$ ( $C_y$ は拗音)	11	
	$C_1V_0C_2$ (V は音韻的無声化)	17	
英語	C	21	1~2 音節
	CC	39	
	CCC	9	
スペイン語	C	18	2 音節
	CC	11	

## 2.2 刺激音の作成

上智大学音声学研究室スタジオにて、上記の語を単独で発話したものを録音した (DAT: Sony TCD-D10、マイク: Primo EMU-4535 II)。インフォーマントは、それぞれの母語話者男性各 1 名。

録音した音声は、16kHz にダウンサンプリングした (Creative Optical Digital I/O2) 後、DC 補正・雑音除去を行った (Syntrillium Cool Edit 96)。その後、22 次の LPC によって残差信号を求めた (フレームは、512-point, 75%-overlap、ハミング窓使用)。各フレームのパワーは原音声と等しくなるよう調節し、また、約 -6dB/oct の傾斜がかかるようにした。

## 2.3 聴取実験

上記で作成した刺激音を、それぞれの言語の母語話者に聞いてもらい、単語の書き取りをしてもらった。刺激音の提示順序はランダムで、繰り返し回数は 2 回とした。実験は、CALL 教室で PC 上のプログラムを使用して行った。被験者がマウスでクリックをすると音声ヘッドフォンから聞こえ、それを回答用紙に書き取ってもらった。一度提示された刺激音を聞き直すことはできないようにした。

聴取実験の被験者は、日本語母語話者 15 名 (男性 4 名、女性 11 名、19-22 才)、英語母語話者 4 名 (男性 2 名、女性 2 名、21-23 才)、スペイン語母語話者 1 名 (女性、24 才) である。

### 3. 結果と分析

#### 3.1 子音連鎖の識別

子音連鎖を正しく聞き取れた率は、図 1 に示す通りである。図 1 は、(1)子音が原音声に含まれているより少なく聞かれたもの（例、C Cが C と聞かれた場合）、(2)子音が多く聞かれたもの、(3)1 つ以上の子音が誤って聞かれたもの、(4)すべての子音が正しく聞かれたものの割合を示している。これを見ると、英語の C C と C C C において、子音数を少なく聞き取ってしまうという誤りが多いことがわかる。

子音を少なく、あるいは多く聞かれてしまった場合は、どの子音がどの子音に聞かれたかという対応を取ることができない。本稿では、子音連鎖という音素配列的制約の中での子音の認識を調べるのが目的なので、次節以下では子音数が増減した場合を除外して分析をする。

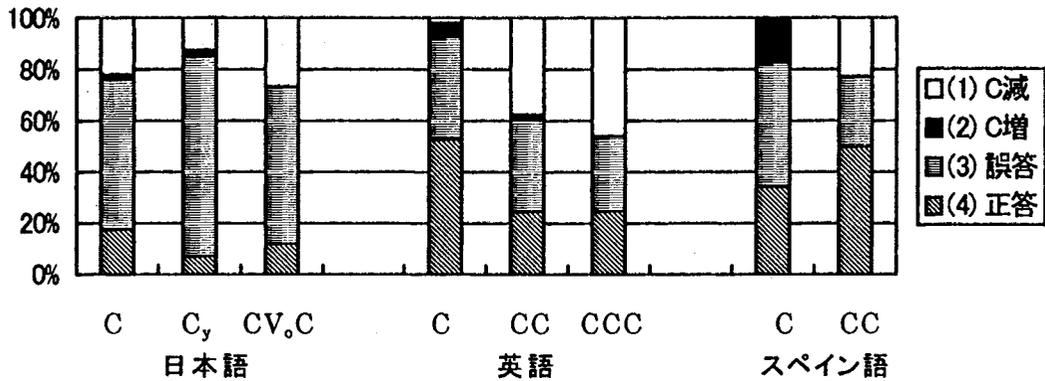


図 1 子音連鎖の識別率

#### 3.2 子音連鎖と子音の識別

子音数が増減した場合を除いた、子音連鎖と其中的各子音の識別率を図 2 に示す。日本語の C と C<sub>y</sub>、英語・スペイン語の C では、子音連鎖と子音の識別率は同じである。

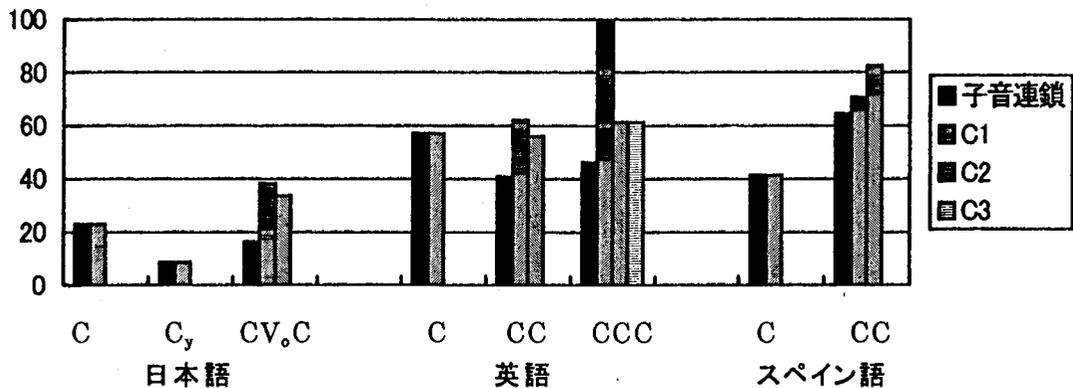


図 2 子音連鎖と子音の識別率

各言語で可能な子音連鎖の数とその連鎖中で可能な子音の数を、本実験で用いた定義に従って表 2 に示す。可能な数が少ないほど聴取した時の判断の選択肢が少ないので、可能な数が少ない子音連鎖ほど識別率が高いことが予想される。図 2 を見ると、スペイン語は明からにこの傾向を示している。日本語の C と CV<sub>0</sub>C は、可能な数にあまり差がなく識別率にもあまり差がないとも見れる。しかし、英語で可能な数が圧倒的に少ない CCC は、もっと識別率が高くなるはずである。そこで、英語の CCC の回答パターンを調べてみると、ほとんどが日常の出現頻度が高いと思われる /spr, str/ と聞かれていた。つまり /spr, str/ 以外の刺激音は誤って聞かれており、その結果識別率が下がったと考えられる。

一方、子音連鎖内での各子音の識別率を見てみると、おおよそ可能な数が少ないほど識別率が高いという傾向が見られる。例えば、図 2 で日本語の CV<sub>0</sub>C の C<sub>1</sub> と C<sub>2</sub> を見ると、C<sub>1</sub> 位置の方が可能な数が少ないので C<sub>2</sub> よりも識別率が高い。英語の CCC やスペイン語の CC も同様の傾向を示していると言えよう。

上記の結果より、子音連鎖および子音の識別率に可能な子音連鎖数・子音数という音素配列的な制約の影響が見られると言える。ただし、出現頻度等そのほかの要因も影響している。

表 2 各言語における可能な子音連鎖の数と子音の数

		子音連鎖	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>
日本語	C/C <sub>v</sub>	26	26	—	—
	CV <sub>0</sub> C	27	3	9	—
英語	C	22	22	—	—
	CC	39	15	10	—
	CCC	9	1	3	4
スペイン語	C	18	18	—	—
	CC	11	12	2	—

### 3.3 調音方法・調音位置・有声／無声の認識

図 3 に、各子音の調音方法・調音位置・有声／無声が正しく認識された率を示す。調音方法の認識率が調音位置より高いことがわかる。また、有声／無声は非常に認識率が高い。(調音方法ごとに認識率は異なる傾向があり、これが前節で述べた子音の識別率に影響している可能性はある。)

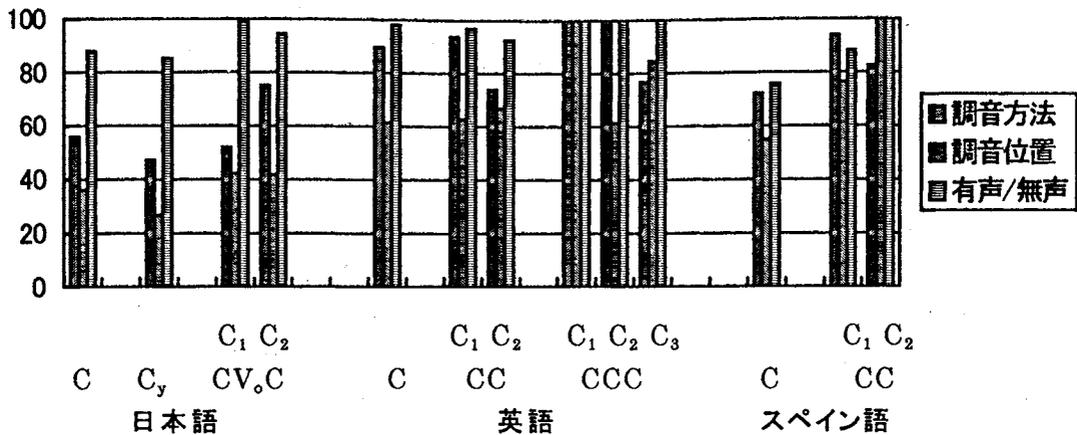


図 3 調音方法・調音位置・有声/無声の認識率

### 3.4 Major Class Features による分析

[consonantal]と[sonorant]は、表 3 に示す通り子音を分類する音韻的素性である(Kenstowicz, 1994)。音韻的な定義では、表中、下に行くほどソノリティが高くなる。英語の/r/は、本来流音に分類されるべきだが、接近音なので/j, w/と同様渡り音に分類した。

表 3 Major Class Features

子音の類	[consonantal]	[sonorant]	ソノリティ
阻害音	+	-	低
鼻音	+	+	:
流音	+	+	:
渡り音	-	+	高

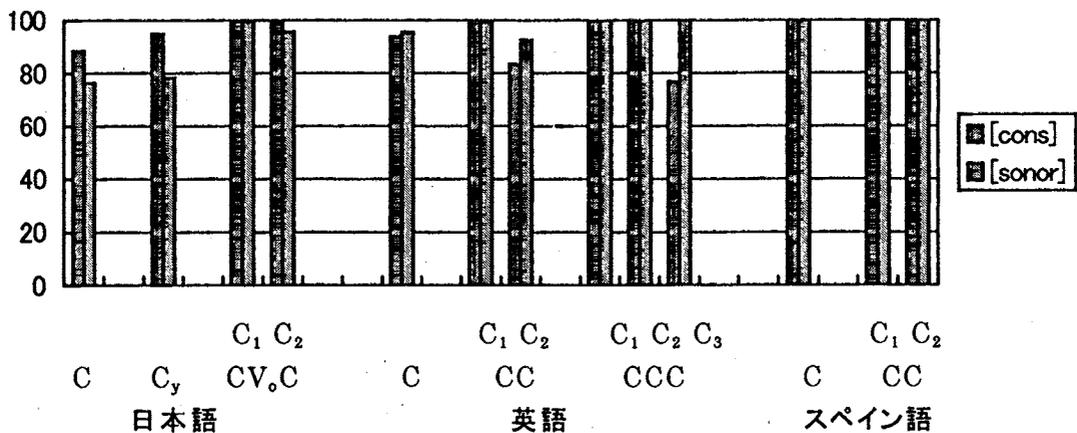


図 4 Major Class Features の認識率

この 2 つの音韻的素性がどれだけ正確に認識されているかを、図 4 に示す。全体としてかなり高い確率で認識できていることがわかる。

ソノリティ情報、すなわちシラブル・リズムに関わるような情報はかなり保たれていると言える。(日本語を用いた予備実験の結果、シラブル数やモーラ数が原音声に含まれているより多く聴取されたり少なく聴取されたりすることはほとんどなかった。)

#### 4. 結論

インテンシティー情報を付加した残差信号による子音の識別率には、音素配列的制約が影響していることがわかった。ただし、出現頻度や調音方法等も関与している可能性がある。言語間の差等、不明な点も多い。また、この残差信号中には、分節素性的情報(子音の調音方法)とも超分節素性的情報(シラブル・リズム)とも見なせるものが含まれている。

#### 参考文献

- 猪塚元・猪塚恵美子(1993)『日本語の音声入門：解説と演習』東京：バベル・プレス。
- 出口厚実(1997)『スペイン語学入門』東京：大学書林。
- 前川喜久雄(1988)「母音の無声化」『講座日本語と日本語教育 2 日本語の音声・音韻(上)』杉藤美代子編。東京：明治書院。135-153。
- 御園和夫(1995)『英語音声学研究：理論と応用』東京：和広出版。
- Arai, Takayuki, Misha Pavel, Hynek Hermansky, and Carlos Avendano (1999) "Syllable Intelligibility for Temporally Filtered LPC Cepstral Trajectories," *J. Acoust. Soc. Am.* 105: 2783-2791.
- Canellada, María Josefa, y John Kuhlmann Madsen (1987) *Pronunciación del Español*. Madrid: Editorial Castalia.
- Kenstowics, Michael (1994) *Phonology in Generative Grammar*. Cambridge, MA: Blackwell.
- Komatsu, Masahiko, Won Tokuma, Shinichi Tokuma, and Takayuki Arai (2000) "The Effect of Reduced Spectral Information on Japanese Consonant Perception: Comparison between L1 and L2 Listeners." Submitted to International Conference on Spoken Language Processing 2000.
- Mori, K., N. Toba, T. Harada, T. Arai, M. Komatsu, M. Aoyagi, and Y. Murahara (1999) "Human Language Identification with Reduced Spectral Information." *Proceedings of Eurospeech '99*. 391-394.
- Quilis, Antonio (1988) *Fonética Acústica de la Lengua Española*. Madrid: Gredos.
- Tsuchida, Ayako (1997) *Phonetics and Phonology of Japanese Vowel Devoicing*. Doctoral dissertation, Cornell University, Ithaca, NY. (Available from UMI, Ann Arbor, MI)