

聴覚障害児の語音識別における Temporal Masking — 単語中の Closure duration の及ぼす影響 —

太田 富雄* 吉野 公善

本研究は、感音難聴児11名を対象に、単語という音声環境を構成している音素の識別にとって CD (closure duration) が及ぼす影響を temporal masking の観点から検討を加えたものである。得られた結果より、以下の点が明らかとなった。

1) CD の値が増加するにつれ識別率は増加する傾向にあったが、CD を1/4以下に短縮しても、また逆に1.5倍以上に伸長しても識別率は変化せず、CD の増加による聞き取りの改善には有効な範囲がある。2) backward masking よりも forward masking の方が識別に及ぼす効果が大きい。3) 周波数領域における音素の主要エネルギー分布でみると、downward よりも upward の方がマスキング効果が大きい。

キーワード：聴覚障害 語音識別 継時マスキング フォワードマスキング バックワードマスキング

I はじめに

音声のように継時的に呈示される情報の知覚を調べるためには、聴覚系における知覚の時間的過程を解明しなければならない。

音声は時間とともに音圧、周波数が変化していくものであり、瞬時に変化する子音の過渡的情報、特に第2ホルマント周波数変移の知覚は重要視されてきた。従来より、連続して入力する刺激音が持つ周波数情報の処理については、前後の音が相互に影響を及ぼすことが示されており、temporal masking の観点からの研究が数多くなされてきた(江端ら, 1980; 北守, 1980, 1981, 1983; Massaro, 1971, 1975)。

感音難聴者の語音識別が劣る原因の1つとして、この temporal masking の影響が考えられており、マスキングの影響を受ける時間間隔は、健聴者よりも長いこと(Elliott, 1975)、第1ホルマントの開始時間を遅らせることで第2ホルマントの弁別閾値が上昇したこと(Danaher and Pickett, 1975)、CV音節の識別においては、後続母音の「音圧の減衰×持続時間の短縮」により識別率が上昇

したこと(太田・吉野1986)などが指摘されている。

ところで、単語を識別する際にも、単語を構成している音素が隣接する環境音によって影響を受けることが推察され、従来の研究から、閉鎖子音の知覚にとって CD (closure duration) は約20ms以上必要であることや解放破裂音圧と CD との間に代償関係(trading relation)のあることが報告されている(Dorman et al., 1979; Fitch et al., 1980; Marcus, 1978; Repp, 1984)。

日本語の単語の場合、V, CV音節が基本の単位となっており、音節と音節の間に CD が存在する。感音難聴者の識別においては、前述のマスキングの影響が現れることが予想され、CD の及ぼす影響を検討することにより感音難聴者における知覚の時間的過程の一端が明らかにされるだろう。

II 目的

本研究は、感音難聴児を対象に、単語という音声環境を構成している音素の識別における音素間の CD の効果を temporal masking の観点から明らかにすることを目的とする。

III 方法

1. 被験児：公立小学校・中学校難聴学級に在

* 心身障害学研究所

Table 1. 被験者の聴力レベル

被験者	年齢	Hearing level (JIS 1982)						SDS*
		250	500	1000	2000	4000	8000(Hz)	
1	10	40	50	60	60	50	45	85/80
2	10	60	60	65	70	65	55	95/95
3	13	45	50	60	65	60	65	85/100
4	14	70	90	85	70	60	70	90/120
5	11	80	75	90	100	90	85	70/90
6	13	25	25	45	60	70	65	90/90
7	13	55	75	75	80	90	90	75/115
8	13	35	25	25	90	90	90	70/60
9	14	70	70	70	70	45	45	100/110
10	14	70	70	60	70	85	95	85/110
11	15	65	65	70	65	55	60	85/110

*SDS=語音弁別能 (日本オーディオロジー学会67式語表)

籍する感音難聴児11名。Table 1に被験児の聴力レベルを示す。被験児はすべて日本オーディオロジー学会67語表中の/s/及び/u/が正答であった者たちである。

2. 刺激：東京方言を話す成人女性によって発せられた「すてき [steki]」, 「うてき [uteki]」の音声を、12ビット量子化、20KHz サンプリング方式でA/D変換を行い、音声切り出し編集装置 (DANAC 7000) に取り込んだ。/s/と/u/の持続時間 (102.4ms)、音圧 (RMS) を同一にし、後続する/te/の音圧 (RMS) とも同一とした。「すてき [steki]」における/s/と/t/、「うてき [uteki]」における/u/と/t/との間の closure duration (CD) を変数としたが、オリジナルの発声4回の平均CD値が78msであったので80msを標準とし、CDを以下の条件となるよう編集した。

N: 標準 (80ms) L₁: 1.5倍 (120ms)L₂: 2.0倍 (160ms) S₁: 1/2倍 (40ms)S₂: 1/4倍 (20ms) S₃: 0 ms

Fig. 1に刺激の構造を示す。これらの刺激語をランダムに配列し、呈示時間、音圧が一定となる

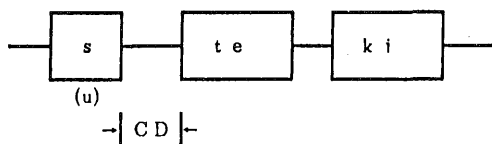


Fig. 1. 刺激の呈示パターン

ように音声切り出し編集装置で制御してテープレコーダ (SONY TC-8750-2) に録音した。

3. 手続き：刺激系列は、防音室内で受話器を用いて被験児の良聴耳に至適レベルで呈示した。

被験児は、「すてき (素敵)」、「せき (席)」、「てき (敵)」、「えき (駅)」、「うてき (雨滴)」、「うえき (植木)」の6単語 (文字カード) の中から強制選択法により pointing で応答した。

IV 結 果

1. 「すてき」, 「うてき」の識別率

Fig. 2は、正答である「すてき」, 「うてき」の識別率を表したものである。識別率を角変換した後、音素(2)×CD(6)の分散分析を行ったところ、

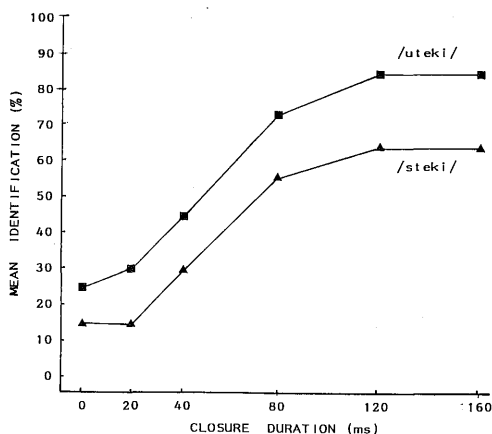


Fig. 2. /steki/, /uteki/の識別率

音素 ($p < 0.01$), CD ($p < 0.01$) の主効果が有意であった。さらに, CD について単純効果を検定したところ, $S_2(20\text{ms})-S_3(0\text{ms})$, $L_1(120\text{ms})-L_2(160\text{ms})$ 間で有意差がみられず, CD を $1/4$ 以下に短縮しても, また逆に 1.5 倍以上に伸長しても識別率は変化しなかった。識別率が著しく変化したのは $N(80\text{ms})-S_1(40\text{ms})$ 間であった。CD の増加に伴い識別率が上昇する過程を logistic 曲線で近似させてみたものが Fig. 3 である。

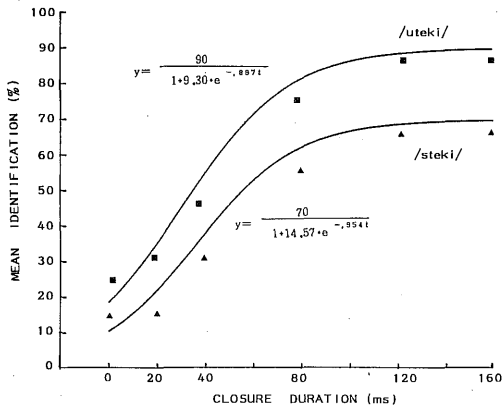


Fig. 3. /steki/, /uteki/ の logistic 曲線

2. Forward masking について

Fig. 4 は, 刺激語「すてき」に対する「せき」の識別率および刺激語「うてき」に対する「うえき」の識別率を表したものである。この図から, 「すてき」の/s/および「うてき」の/u/が, それぞれの単語中において/t/をマスキング (forward masking) したことがみとめられる。音素(2)×CD(6)の分散分析を行ったところ, 音素 ($p < 0.01$), CD ($p < 0.01$) の主効果が有意であった。このことは, 音素に関しては/s/よりも/u/の方がマスキング効果は大きく, CD に関しては CD が長くなるにつれてマスキング効果は弱まることを示している。

3. Backward masking について

Fig. 5 は, 刺激語「すてき」に対する「てき」の識別率および刺激語「うてき」に対する「てき」の識別率を表したものである。この図から, 「すてき」の/t/が単語中において/s/をマスキング (backward masking) したことは認められるが, 「うてき」の/t/が単語中において/u/をマスキ

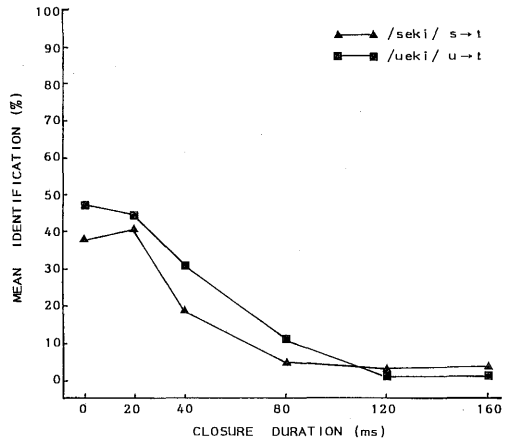


Fig. 4. 音素による forward masking

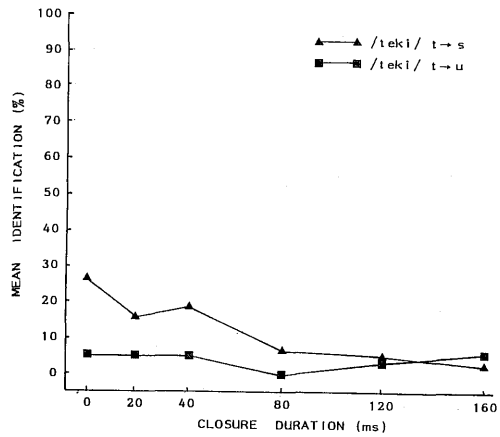


Fig. 5. 音素による backward masking

グ (backward masking) したことはみとめ難い。音素(2)×CD(6)の分散分析を行ったところ, 音素 ($p < 0.01$), CD ($p < 0.01$) の主効果が有意であった。

V 考 察

「すてき」, 「うてき」の識別率をみると, CD の増加に伴う識別率の上昇, 音素による識別率の違いがみられた。Fig. 3 の曲線をみると, CD を増加させても識別率は一定値に収束してしまい, CD 増加させることで難聴児の聴取を改善させようとしても, CD 増加の程度には有効な範囲があることを示唆している。北守(1981, 1983)は, 短音の周波数弁別の時間的増大過程について指数関数モデルを用いて検討している。両モデル

とも一定値に収束してしまうことは同じであるが、時間を独立変数とするとき、処理能力の上昇する傾向が異なる。つまり、変数の効果が現れるのが難聴児は最初緩やかであるのに対し健聴者は急峻となることを意味するが、課題、処理の次元等が異なり今後さらに検討する必要がある。

CDが識別率に及ぼす影響が顕著になってくるのは20ms以降であった。Marcus (1978)は、語中のSI (SI: silent interval)を操作し、30~45msのSIが識別にとって重要であることを示し、Huggins (1972)が生理的限界から予測した40msの最小知覚CDと一致することを指摘した。また、Repp (1984)は、[s]ノイズに続く閉鎖子音の知覚を調べ、[s]ノイズと定常な合成母音間のSIが約20msよりも長いと閉鎖子音知覚を生じ、閉鎖様式の他の知覚手掛りが信号中に存在していてもSIは閉鎖子音知覚にとって必要であると述べている。北守 (1981)は、聴覚系での周波数情報処理システムは時定数50~60ms程度の1次遅れシステムとして近似することができるとしている。これらの値には差がみられるが、刺激の種類や性質、実験手続き、被験者等の違いがあるため一概に比較することはできず、今後もデータの集積が必要であろう。

一方、音素による識別率の違いについてみると、刺激語の/s/と/u/は音圧、持続時間共に同一にしてあり、波形構造よりも周波数構造が問題となる。両者の周波数構造における音響的特徴をみると、/s/は高周波数帯域にエネルギーが分布しているのに対し、/u/はホルマント構造が明確であり、両者の音響構造の違いが識別率に反映することが示唆される。

forward masking 及び backward masking においては、CDの増加に伴うマスク効果の減少、音素によるマスク効果の違いがみられた。forward masking 及び backward masking においても刺激語の周波数構造が問題となる。

forward masking においては、/s/よりも/u/の方が/t/に対するマスク効果は大きかった。周波数領域における音響エネルギーの分布について、/s/から/t/を見れば下向き (downward)、/u/から/t/を見れば上向き (upward)となる。つまり、マスクする音からマスクされる音を周波数領域においてみた場合、上向きの方がマスク効果が大きいことがわかる。

backward masking においても、同様に/t/→/s/という上向きの方が/t/→/u/という下向きよりもマスク効果が大きかった。

本実験では、/s/と/u/の音圧を後続する/te/とも同一にした。[seki]及び[ueki]という反応は、/s/及び/u/が/t/を forward masking した結果得られるのと同様に、/e/が/t/を backward masking した結果得られたものと解釈することもできる。太田・吉野 (1986)は、日本語 CV 音節の母音が子音の識別に及ぼす backward masking 効果を調べたが、/te/の場合、先行子音/t/の識別手掛りとなる第2ホルマント周波数に対して後続母音/e/の第2ホルマントが高く、周波数領域において/e/から/t/をみた場合、下向きとなり backward masking 効果が小さかったことを示している。本実験において、もし/e/が/t/を backward masking したのなら、CDの値にかかわらず[seki]、[ueki]の識別率が一定となっている筈であるが、CDの増加に伴い識別率が低下しており、この解釈は成り立たないことになる。

江端ら (1980)は、短音の周波数弁別に及ぼすマスク効果は先行音よりも後続音の方が大であると指摘しているが、本実験の識別率でみる限り、forward masking の影響の方が大きかった。江端らは、先行音がマスクするマイナスの効果よりも、先行音とテスト音のピッチ差の情報が周波数弁別の手掛りを与えるというプラスの効果の方が大きいことによると解釈している。弁別 (discrimination) と識別 (identification) とでは処理の次元が異なり、また、刺激の種類、被験者が違うこともあって単純に比較することはできないが、マスクのパラダイムにおいてもプラスの効果があるということは、刺激の性質や呈示方法を検討することで難聴児の聞こえの改善にもつながるものと推測される。

先に述べたように、本実験では/s/及び/u/の音圧を/te/と同一にしたが、/t/の音圧、持続時間等が識別に及ぼす影響も問題になる。Repp (1984)は、自然音声の歯茎音の解放破裂はSIが存在しない時間閉鎖子音の知覚にとって充分な手掛りとなることを示したが、その手掛りは解放破裂の最初の数ms、すなわち、急激な音響エネルギー増加の絶対値と傾斜 (立ち上がり)に存在することを示唆している。また、SIの役割はそれらの特徴の聴覚的呈示に対する先行摩擦ノイズの forward

masking 効果を防ぐこと及び聞き手が重要な開始特性 (onset properties) に注意を向けることができるようにすることだと指摘している。これらの点についての検討は、今後の課題として残る。

文 献

- 1) Danaher, E.M. and Pickett, J.M. (1975): Some masking effects produced by low-frequency vowel formants in persons with sensorineural hearing loss. *J. Speech Hear. Res.*, 18, 261—271.
- 2) Dorman, M.F., Raphael, L.J., and Liberman, A.M. (1979): Some experiments on the sound of silence in phonetic perception. *J. Acoust. Soc. Am.*, 65, 1518—1532.
- 3) 江端正直・山内修・曾根敏夫・二村忠元 (1980): 短音の周波数弁別に及ぼす先行音および後続音の影響. *日本音響学会誌*, 36, 232—237.
- 4) Elliott, L.L. (1975): Temporal and masking phenomena in persons with sensorineural hearing loss. *Audiology*, 14, 336—353.
- 5) Fitch, H.L., Halwes, T., Erickson, D.M., and Liberman, A.M. (1980): Perceptual equivalence of two acoustic cues for stop-consonant manner. *Percept. Psychophys.*, 27, 343—350.
- 6) Huggins, A.W.F. (1972): Just noticeable differences for segment duration in natural speech. *J. Acoust. Soc. Am.*, 51, 1270—1278.
- 7) 北守進(1980): 聴覚系での周波数情報処理能力の時間モデル. 沼津工業高等専門学校研究報告, 16, 37—40.
- 8) 北守進(1981): 短音の周波数弁別における処理時間に関する考察. *日本音響学会誌*, 37, 446—449.
- 9) 北守進(1983): 短音のピッチ認識に及ぼす後続音の効果に関する検討. *日本音響学会誌*, 39, 82—88.
- 10) Marcus, S.M. (1978): Distinguishing “slit” and “split”—an invariant timing cue in speech perception. *Percept. Psychophys.*, 23, 58—60.
- 11) Massaro, D.W. (1971): Effect of masking tone duration on perceptual auditory image. *J. Exp. Psychol.*, 87, 146—148.
- 12) Massaro, D.W. (1975): Backward recognition masking. *J. Acoust. Soc. Am.*, 58, 1059—1065.
- 13) 太田富雄・吉野公喜 (1986): 聴覚障害児の語音知覚における Backward Masking—破裂語音の明瞭度を中心に—. *心身障害学研究*, 10(2), 23—30.
- 14) Repp, B.H. (1984): The role of release bursts in the perception of [s]-stop clusters. *J. Acoust. Soc. Am.*, 75, 1219—1230.

Summary

Temporal Masking Effects on Speech Perception in Hearing-Impaired Children —Effect of Closure Duration in Word—

Tomio Ohta Tomoyoshi Yoshino

Listeners with sensorineural hearing loss have been shown to have poor speech identification. From the data of temporal masking, we might speculate the effects of forward and backward masking on speech perception in persons with sensorineural hearing loss.

In word perception, there is evidence that the durations of the acoustic segments preceding and following the closure can influence stop consonant perception.

The purpose of present study was to clarify the effect of closure duration on word perception in children with sensorineural hearing loss. The subjects were 11 children with moderate to severe sensorineural hearing loss. Good tokens of [steki] and [uteki] produced by a female speaker were served as stimuli. The closure duration (CD) between /s/ (/u/) and /t/ was varied from 0 to 160ms. The rms amplitudes of /s/ and /u/ were equalized with /te/.

The results were as follows. (1) In correct and forward masked identification, the results of the two way analysis of variance indicated the main effects of CD and phoneme. However, we couldn't observe the main effect of CD in backward masked identification. (2) The improvement in correct identification of [steki] and [uteki] can be well fitted with the logistic model. (3) Interference effect of forward masking on speech identification was greater than that of backward masking. (4) From the viewpoint of frequency structures, upward spread of masking reduced the phoneme information.

Key word: hearing-impaired, speech identification, temporal masking forward masking, backward masking