

音声信号の一部を間欠的に削除した場合の 明瞭度について (第2報)

石原 学*・白滝 順**・家入 勝吾**

Intelligibility of Intermittently Eliminated Japanese Speech Wave (2nd Report)

Manabu ISHIHARA, Jun SHIRATAKI and Shogo IEIRI

Abstract

The intelligibility of intermittently eliminated Japanese speech wave is discussed in this paper. In our experiment, a segment of a Japanese speech wave is intermittently removed by varying the erasing duration and the period. Here, the ratios of the erasing duration to period are varied from 1:9 to 9:1 every 10 percent.

Listening tests for evaluating the intermittently eliminated speech wave have been carried out by about 600 males aged from 21 to 25 years. Experimental results show that the intelligibility of the composite tones deteriorates linearly. In the case of the period being short, however, the intelligibility of speech wave is more than 90% even though ninety percent of the speech wave in one period is removed.

1. はじめに

音声信号の持つ特徴については現在までに各種の研究が行われてきている¹⁻⁸⁾。

音声信号は、言語情報として聞きとることが出来て人間対人間のコミュニケーションに役立つものであり、その機能を機械に持たせることにより人間対機械間の有効なコミュニケーションができると思われる。そのため、音声研究に対する各種のアプローチが必要であると考えられる。

これらの応用例として、時間間引き法による音声信号の時間圧縮や⁹⁾、日本語単音節の構造を研究するために語頭を切り取った場合の特徴について報告されている¹⁰⁾。また、時間間引きによる音声合成信号の特徴について検討されている¹¹⁾。さらに、音声信号を時間間引きした場合の聞きとりについても検討されている¹²⁾。この研究は、音声信号を一定間隔の周期ごとに削除し

て明瞭度の特徴を検討したものである。その結果、一周期あたりの削除区間が長くなるに従って、明瞭度が直線的に低下すること、さらに、周期が短い場合には、音声波と削除区間の割合が1:9のとき、すなわち一周期あたりの削除区間が大きい場合でも、明瞭度が90%以上あるという実験結果を得た。そのときの被験者のヘッドホン出力電圧を調べると、単音節の削除区間が長い場合には、より大きな出力電圧を必要とすることを報告している¹²⁾。

本論文では、音声信号をある時間間隔で削除した場合、明瞭度が低下する閾値を求めるために、周期(信号区間+削除区間)を、0.1, 0.143, 0.2, 3.33, 5.0および20 msの6通りに変化させて実験を行った。その結果、部分的に削除された音声波を人間がどのように聞きとるのかについて新しく得られた知見について報告する。

実験の結果より、一周期で音声信号を削除する割合が大きくなるにしたがって明瞭度が低下し、そのときの明瞭度の低下は、一周期の時間長によって異なることを明らかにした。また、明瞭度は、一周期中に含まれる音声信号の割合が50%程度までは直線的に増大

平成2年9月28日受理

* 千葉経済短期大学

** 電気工学科

するが、さらに音声区間が増加すると、傾斜が次第にゆるやかになり、ついには飽和する傾向がみられた。

なお、削除区間を変化させて各条件ごとに明瞭度の標準偏差を計算した結果によると、一周期が0.2 msより短いときには、明瞭度の標準偏差は±6%程度で、個人によるばらつきが少ない。しかし、一周期が3.33 msより長い場合には、明瞭度のばらつきが最大で±20%～±30%程度みられる。したがって、削除区間が長い場合には、個人差による影響が現われることがわかった。

2. 実 験

本実験の回路構成を図1に示す。まず、音声波をD.A.T. (SONY社, DTC-1000ES)より入力し、演算増幅器を利用したバターワース特性を有するバンドパスフィルタ(周波数特性300~5,000 Hz, 48 dB/oct)を通した後、この音声信号を、波形処理できる大きさまで演算増幅器を用いて増幅する。その後、アナログスイッチを利用して音声波を一定の時間間隔で削除する。このアナログスイッチは、FETで構成されている。また、製作したスイッチング・パルス・ジェネレータは、可変周波数範囲1 Hz~2 MHzで、TTLレベルの方形波を発生できる。このように回路を設計・製作することによって、音声波と削除区間が交互に合成される。このときの合成波形を削除音声信号と呼ぶことにする。明瞭度実験では、削除音声信号の一周期は、0.1, 0.143, 0.2, 3.33, 5.0 および 20 msの6通りのうちから選択でき、かつ各周期について音声信号区間と削除区間の割合(分割比)を9:1から1:9まで、10%おきに変化させることができる。

実験に用いた音声資料¹³⁾としては、NHKアナウンサーによる100単音節の発声テープを利用した。被験者は、21歳から25歳までの健康な男性5名、延べ約

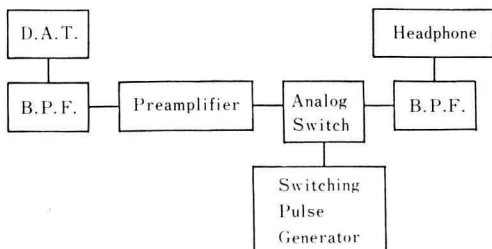


図1. 実験装置のブロック図

600名である。明瞭度の測定は、被験者5名が同時に、ランダムに並び変えられた単音節の削除音声信号をヘッドホン(バイオニア社, SE-M70)で聞きとり、聞きとった音声を解答表に記入する方式で行われた。明瞭度は、一単音節あたり正解のとき1%, 100単音節すべて正解のとき100%, と評価される。

3. 音声波の削除区間と明瞭度の関係

著者らは、すでに周期が大きくて(10.0および12.5 ms)、分割比が小さい場合、および周期が小さくて(50 μs)分割比が大きい場合の明瞭度特性について検討した^{8,12)}。本論文では、さらに削除音声信号の周期を0.1, 0.143, 0.2, 3.33, 5および20 msの6通りについて実験を行った結果について述べる。実験結果を図2に示す。図2で、横軸は(音声区間:削除区間)の分割比ではなくて、次式のようにとってある。

$$\ln \{ (\text{音声区間} / \text{削除区間}) \times 100 \}$$

ここでは、上式で与えられる量を便宜上対数分割比と称する。

図2からわかるように、音声信号の削除区間が短くなるにしたがって単音節音声の明瞭度は高くなっていく。本実験のデータから、周期が0.2 msより小さい場合には80%以上の明瞭度が得られている。80%以上の単音節明瞭度が得られていることは、90%以上の高

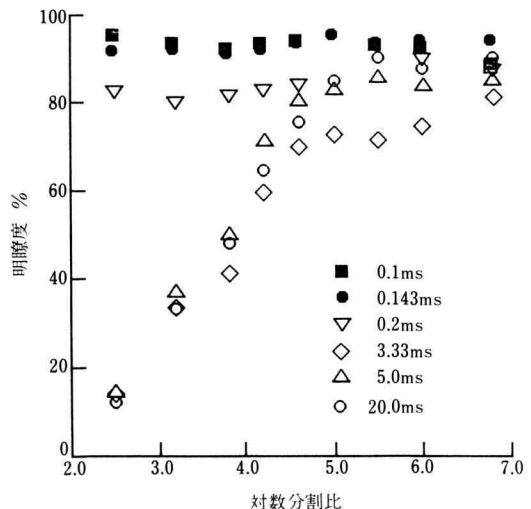


図2. 対数分割比に対する明瞭度

い文章理解度が期待できるといえる。さらに、音声区間：削除区間の割合が(1:9)のときでも、(9:1)の場合のように、90%以上の明瞭度が得られており、分割比による変化がほとんどみられない。このことは、音声信号を削除していき、削除区間が大きい場合でも十分な明瞭度が得られることを示している。

また、削除音声信号の周期が大きい場合については、削除される割合が大きくなるにしたがって、分割比に

対する明瞭度の傾斜が著しい変化を起こす部分は音声信号の少ない場合に移動する傾向がみられた。

さらに、削除区間を変化させて各条件ごとに標準偏差を計算した結果、一周期が0.2msより短いときには、標準偏差は±6%程度で、個人によるばらつきが少ない。しかし、一周期が3.33msより長い場合には、ばらつきが最大で±20%~±30%程度みられる。すなわち、削除区間の長い場合には、個人差による影響がみ

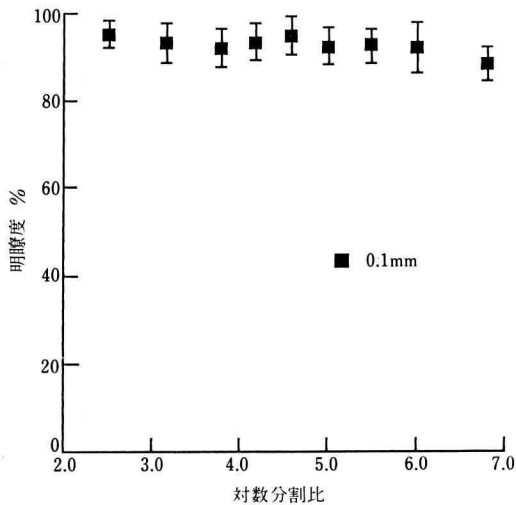


図3. (1) 対数分割比に対する明瞭度 (0.1msの場合)

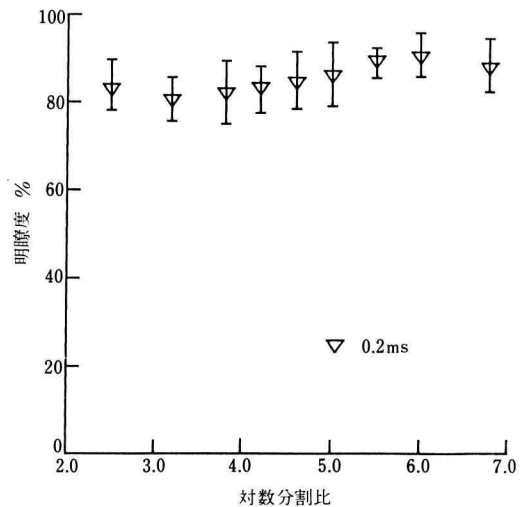


図3. (3) 対数分割比に対する明瞭度 (0.2msの場合)

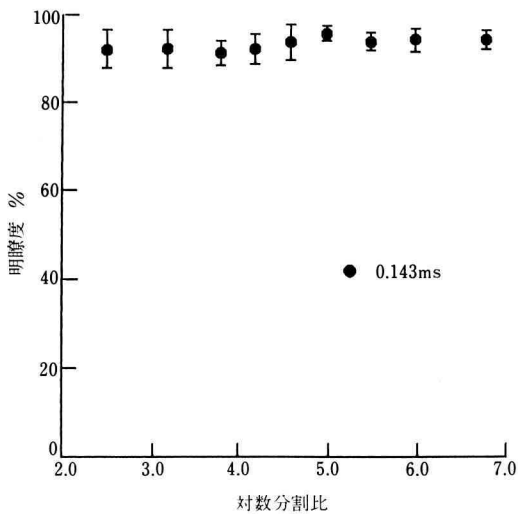


図3. (2) 対数分割比に対する明瞭度 (0.143msの場合)

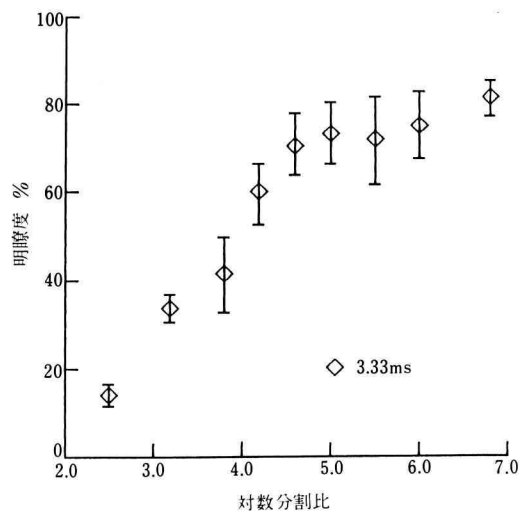


図3. (4) 対数分割比に対する明瞭度 (3.33msの場合)

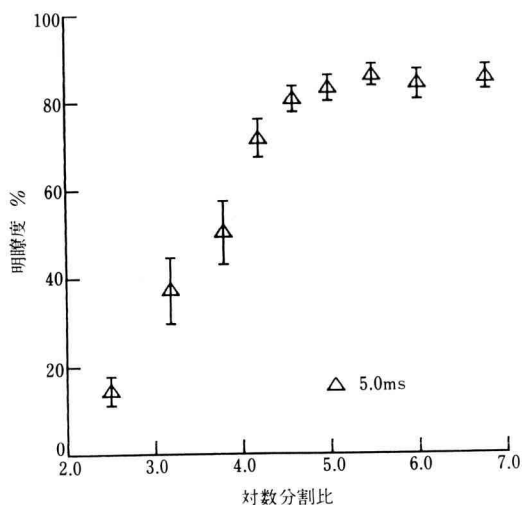


図3. (5) 対数分割比に対する明瞭度 (5.0 ms の場合)

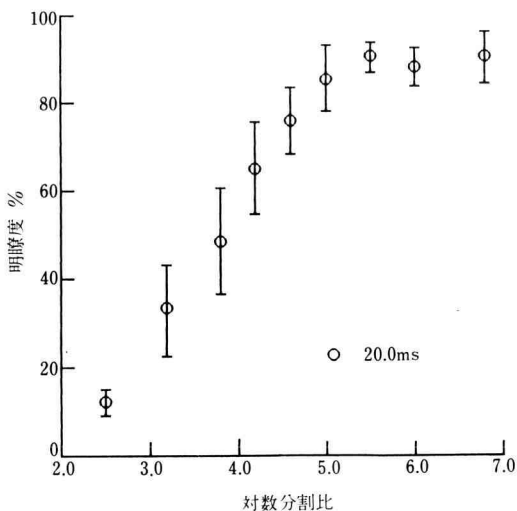


図3. (6) 対数分割比に対する明瞭度 (20.0 ms の場合)

られた。このときの、各削除区間ごとの明瞭度特性を図3(1)~(6)に示す。図3からわかるように、音声信号が多く含まれる場合と、少ない場合の明瞭度特性にはバラツキが少ない。すなわち、明瞭度特性の中心部分よりのほうがバラツキが大きい。このことは、音声信号を認知するとき、音声の含まれる割合によって、各個人の聞きとりやすい単音節音声異なることによる、と考えられる。

また、周期が0.2 msより短く明瞭度が高い場合について、各子音を系統別に分類し、聞きとり方について検討した。その結果、各条件の下で異聴の相違が見られた。そのときの異聴の傾向を表1から表9に示す。表1は、周期が0.1 msで、分割比が(9:1)のときの様子を示す。そのときには破裂音は破擦音に聞き間違える

ことが多い。しかし、削除区間が長くなるにしたがって、破裂音は破裂音に聞き間違えられることが多くなり、また弾音は破裂音に聞き間違えられることが多くなる傾向がみられた。その他の子音については、多少他の子音群に聞き間違えられるが、ほぼ同一の子音群に異聴される傾向がみうけられる。表2は、周期0.1 ms、分割比(1:1)の場合を、表3は、周期0.1 ms、分割比(1:9)の場合を、表4は、周期0.143 ms、分割比(9:1)の場合を、表5は、周期0.143 ms、分割比(1:1)の場合を、表6は、周期0.143 ms、分割比(1:9)の場合を、表7は、周期0.2 ms、分割比(9:1)の場合を、表8は、周期0.2 ms、分割比(1:1)の場合を、また表9は、周期0.2 ms、分割比(1:9)の場合をそれぞれ示す。

表1. 子音の系統別異聴表
周期0.1 ms, 分割比(9:1)の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	4	2	0	1	4	0
摩擦音	0	3	0	0	0	0
破擦音	13	0	1	0	0	0
鼻音	1	0	1	11	0	0
弾音	8	0	0	0	1	0
半母音	0	0	0	0	0	0

表2. 子音の系統別異聴表
周期0.1 ms, 分割比(1:1)の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	3	0	0	0	1	0
摩擦音	0	0	0	0	0	0
破擦音	10	0	0	0	0	0
鼻音	0	0	0	9	0	0
弾音	2	1	0	0	0	0
半母音	0	0	0	0	0	0

表3. 子音の系統別異聴表
周期 0.1 ms, 分割比 (1:9) の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	4	0	0	0	3	0
摩擦音	1	0	0	0	0	0
破擦音	6	0	0	0	0	0
鼻音	0	0	0	7	0	0
弾音	3	0	0	0	0	0
半母音	0	0	0	0	0	0

表4. 子音の系統別異聴表
周期 0.143 ms, 分割比 (9:1) の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	3	0	0	1	0	0
摩擦音	1	0	0	0	0	0
破擦音	7	0	0	0	0	0
鼻音	1	0	0	9	0	0
弾音	6	0	0	0	0	0
半母音	0	0	0	0	0	0

表5. 子音の系統別異聴表
周期 0.143 ms, 分割比 (1:1) の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	5	0	1	0	3	0
摩擦音	1	3	0	0	0	0
破擦音	8	1	0	0	0	0
鼻音	0	0	0	2	0	0
弾音	6	0	0	0	0	0
半母音	0	0	0	0	0	0

表6. 子音の系統別異聴表
周期 0.143 ms, 分割比 (1:9) の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	7	1	1	0	2	0
摩擦音	5	4	0	0	0	0
破擦音	7	1	0	0	0	0
鼻音	0	0	0	9	0	0
弾音	1	0	0	0	0	0
半母音	0	0	0	0	0	0

表7. 子音の系統別異聴表
周期 0.2 ms, 分割比 (9:1) の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	3	0	1	0	6	0
摩擦音	2	2	0	1	0	0
破擦音	10	0	0	0	0	0
鼻音	0	0	0	15	0	0
弾音	8	0	0	0	1	0
半母音	0	0	0	0	0	0

表8. 子音の系統別異聴表
周期 0.2 ms, 分割比 (1:1) の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	9	1	2	0	12	0
摩擦音	3	8	0	0	0	0
破擦音	20	1	1	0	0	0
鼻音	0	0	0	11	0	0
弾音	5	0	0	0	1	0
半母音	0	0	0	0	0	0

表9. 子音の系統別異聴表
周期 0.2 ms, 分割比 (1:9) の場合

	破裂音	摩擦音	破擦音	鼻音	弾音	半母音
破裂音	14	0	2	0	12	0
摩擦音	1	10	1	0	0	0
破擦音	14	0	2	0	0	0
鼻音	0	0	0	16	0	0
弾音	9	0	0	0	0	0
半母音	0	0	0	0	0	0

4. む す び

実験データに基づいて、削除音声信号を受聴した場合の明瞭度のもつ特徴について検討した。周期が長い削除音声信号を聞きとったときの明瞭度は、分割比が小さくなるに従って、すなわち音声波の削除区間が長くなるに従って明瞭度がほぼ直線的に低下していることを実験より示した。また、周期が短いときには、分割比が例えば1:9のように小さいときでも、明瞭度が92~98.5%と高いことを実験から明らかにした。また、周期が0.2 msより短い場合には、80%以上の明瞭度

が得られている。80%以上の単音節明瞭度が得られていることは、90%以上の高い文章理解度が期待できるといえる¹⁾。さらに、音声区間:削除区間の割合が(1:9)のときでも、(9:1)の場合のように90%以上の明瞭度が得られ、分割比による明瞭度変化がほとんどみられない。このことは、周期が小さいときには、音声の削除区間が大きい場合、すなわち分割比が小さくても十分な明瞭度が得られることを示している。

また、周期が長いときには、削除される時間間隔が大きくなるにしたがって、明瞭度の傾斜が著しく変化し始める部分は、音声信号の少ない場合に移動する傾向がみられた。

これらのことは、音響工学や人間工学およびヒューマンインタフェースの上で極めて重要と考えられる。

文 献

- 1) 三浦種敏：“日本語に対する定量的な伝送品質と伝送特性との関係”，電電公社通研実報第3巻第4号(1954)。
- 2) Licklider, J.C.R and Pollack, I: “Effects of differentiation, integration, and infinite peak clipping upon the Intelligibility of speech”, J. Acoust. Soc. Am., 20, pp. 42-51 (1948).
- 3) Russell J. Niederjohn, Michael W. Kruts and Bruce M. Brown: “An Experimental Investigation of the Perceptual Effects of Altering the Zero-Crossings of a Speech Signal”, IEEE Tran. on Acoustics, Speech, and Signal Processing, Vol. ASSP-35, No. 5, pp. 618-625 (1987).
- 4) 重永 実：“日本語音声の零交差波時系列 I. 測定方法および母音, II. 無声子音”，音響学会誌, 17(1), pp. 77-92 (1961).
- 5) 石原 学, 白滝 順：“音声波のスライスレベルと明瞭度の関係について”，人間工学, 23(4), pp. 241-245 (1987).
- 6) 松岡教栄, 城戸健一：“音声スペクトルのローカルピークの静特性のもつ音韻情報に関する検討”，音響学会誌, 32(1), pp. 12-23 (1976).
- 7) 例えば, 樋渡涓二編著：“視聴覚情報概論”, pp. 115-121, 昭晃堂 (1987).
- 8) 石原 学, 白滝 順, 家入勝吾, 志方 泰：“単音節音声に混入されたノイズと明瞭度の関係について”，計測自動制御学会ヒューマンインタフェース部会, 第4回ヒューマン・インターフェース・シンポジウム論文集, pp. 321-324 (1988).
- 9) 加藤誠巳, 藤原ひろみ, 佐藤恵臣：“時間間引きによる音声信号の時間軸圧縮に関する基礎検討”，情報処理学会平成元年度前期全国大会, 3K-8, pp. 680-681 (1989).
- 10) 桑原尚夫, 境 久雄：“連続音声中の切り出し母音および音節の音韻知覚”，音響学会誌, 28(5), pp. 225-234 (1972).
- 11) 片桐 滋, 東倉洋一, 古井貞熙：“単音節知覚における時間情報の役割”，音響学会, 42, 2, pp. 97-105 (1986).
- 12) 石原 学, 白滝 順, 家入勝吾, 志方 泰：“音声信号の一部を削除した場合の明瞭度に関する実験的研究”，機械学会論文集C編, 56, 529, p. 92-95 (1990).
- 13) 日本音響学会編：“明瞭度試験法の基準”および“明瞭度試験用テープ”(昭和33年)。