

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

石原 学・白滝 順・家入 勝吾・志方 泰*

Intelligibility for Intermittently Eliminated Japanese Speech Wave

Manabu ISHIHARA, Jun SHIRATAKI, Shogo IEIRI
and Yutaka SHIKATA

Abstract

The intelligibility for intermittently eliminated Japanese speech wave is discussed in this paper.

In our experiment, a segment of Japanese speech wave is intermittently removed by varying the erasing duration and the period. Here, ratios of the erasing duration to period are varied from 10:90 to 90:10 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 for the composite tones deteriorates linearly. In case of the period being short, however, the intelligibility of speech wave is more than 90% even though ninety percent of speech wave in one period is removed.

1. はじめに

音声波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⁾。音声波自身は、波形処理された場合でもある程度までは十分に聞きとることができる²⁻⁴⁾。音声を聞きとるときには音声波の持つスペクトルのローカルピーク値の重要性が指摘されている⁵⁾。また、音声波にノイズを加えたときの聞きとりに対する性質についても、マスキング等の研究として聴覚との関係から報告されている^{6,7)}。

また日本語単音節の構造としては語頭を切り取った場合の特徴について報告されている⁸⁾。音声波は、言語情報として聞きとることができ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、また、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。そのためには、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。

応用例の一つとして、時間間引き法による音声信号の時間圧縮が考えられている⁹⁾が、音声波を時間間引きした場合の特徴についてはあまり検討されていないようである。そこで、われわれは、単音節音声波の一部を一定の時間間隔毎に削除し、またその削除された間隔を変化させたときに、人間がその信号をどのように聞きとるかについて実験を行った。

単音節音声を時間的に分割し、この分割された部分を削除し、その削除された部分を含む音声波を被験者に聞きとらせた。そのとき、音声波と削除区間の割合を1:9(音声波:削除区間)から9:1まで10%毎に変化させて実験を行った。その結果、一周期あたりの削除区間が長いときには、明瞭度が直線的に低下することがわかった。さらに、一周期あたりの削除区間の時間間隔が短い場合については、音声波と削除区間の割合が1:9(音声波:削除区間)と全体の削除区間が大きい場合でも明瞭度が90%以上あることを実験から求めた。そのときの、被験者のヘッドホン出力を調べると削除区間の長い単音節がより大きな出力を必要とす

平成元年 10月2日受理

* Department of Electrical Engineering, Faculty of Science and Engineering, Meisei University.

ることがわかった。

2. 実 験

本実験の回路構成を Fig.1 に示す。まず、音声波を D.A.T. より入力し、バンドパスフィルタ (300~5,000 Hz, 48 dB/oct) を通した後、この音声信号を、波形処理できる範囲まで増幅する。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。

そして、音声波と削除区間が交互に合成されるように回路上設計されている。この2つを合成することによって音声信号として出力して、その音声信号を被験者が聞きとる。本実験では、時間間隔を一周期あたり長い削除区間とするものは 10.0 ms および 12.5 ms とし、一周期あたり短い削除区間とするものは 50 μ s とし分割削除を行うことにした。このようにして、周期ごとに音声波と削除区間が出力される。したがって音声波が削除された区間と一緒に出力する場合の割合は周期が長い場合と同じように出力される。音声と削除区間を合成する比率は、10% の割合で 0~100% まで可変できる。

実験に用いた音声資料¹⁰⁾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、被験者は 21 歳から 25 歳までの健康な男性 5 名で、延べ 600 名である。明瞭度の測定は、ランダムに並び変えられた単音節を被験者が聞きとり、被験者が解答表に記入する方式で行った。一単音節あたり 1% として、100 単音節すべて正解のとき 100% である。Fig. 2(1) は、分割削除された音声波の概形を示している。また、Fig. 2(2) は正弦波と削除区間の関係の例を示す。分割削除される時間の間隔は 0.1 ms とした。Fig. 2(3) は、母音 /a/ を一定間隔の間削除したときの波形である。なお、Fig. 2(2) および Fig. 2(3) の場合、音声波と削除区間の比は、いずれも 1:1 である。

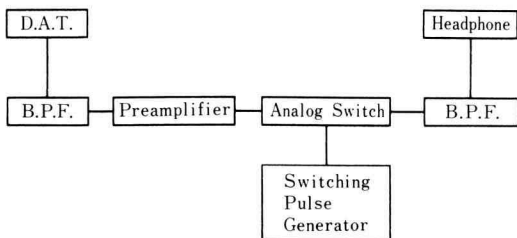


Fig.1 Block Diagram of System

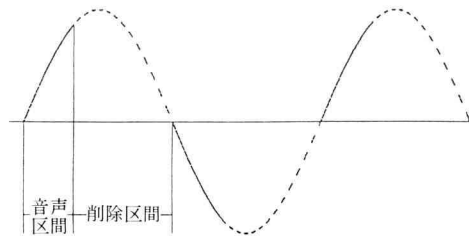


Fig. 2(1) Intermittently Eliminated Japanese Speech

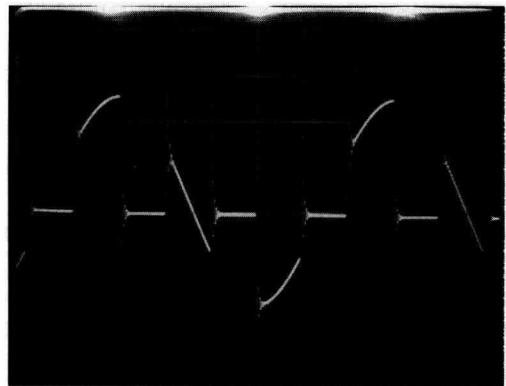


Fig. 2(2) Intermittently Eliminated Sine wave
(X-axis: 50 μ s/DIV, Y-axis: 1 V/DIV)

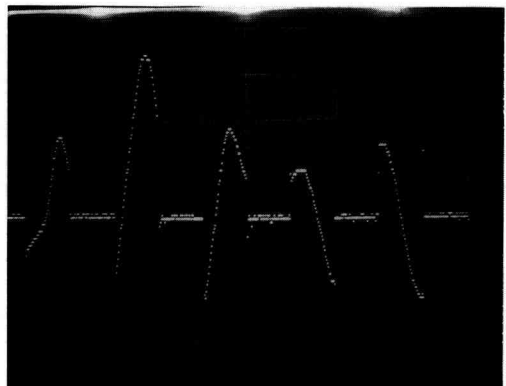


Fig. 2(3) Intermittently Eliminated Vowel /a/

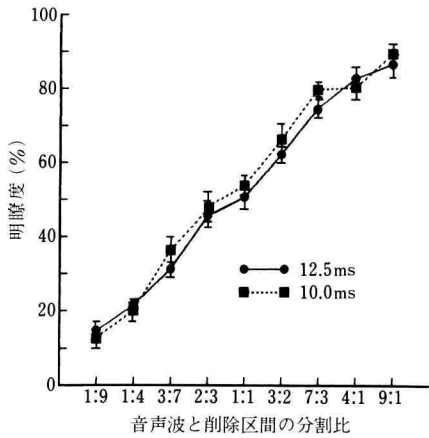


Fig. 3 Intelligibility against Ratios of Speech to Erasing Durations

3. 結果および考察

3.1 音声波の削除区間と明瞭度の関係について

～一周期の削除区間が長い場合～

本実験で削除する時間間隔は、12.5 ms および 10.0 ms の2通りとして、さらに、音声波と削除区間の割合を1:9 から9:1まで(音声波:削除区間)について10% ごとに変化させて実験を行った。音声波と削除区間をそれぞれ9通りに変化させたときの分割比と明瞭度の関係をFig. 3に示す。ここで、音声波:削除区間の割合は1:9, 1:4, 3:7, 2:3, 1:1, 3:2, 7:3, 4:1, 9:1の9通りである。Fig. 3からわかるように、削除区間の占める割合が大きくなるに従って明瞭度が低下する。これは音声波の削除区間が大きくなるにしたがって音声情報が減少するための影響と考えられ、経験的にも理解できる。そのときの明瞭度は、ほぼ直線的に低下する傾向にある。また、削除区間が10.0 msの場合には12.5 msのときよりも明瞭度の低下が大きいの。また、削除間隔が1:4よりも大きい場合と、4:1よりも小さい場合には明瞭度の差がはっきりしない。これは削除される部分の値が聴覚に対して影響を及ぼすためと思われる。

3.2 音声波の削除区間と明瞭度の関係について

～一周期の削除区間が短い場合～

音声波の削除区間の間隔が短い場合の実験結果をFig. 4に示す。

この実験では、音声波を削除する時間間隔を50 μs

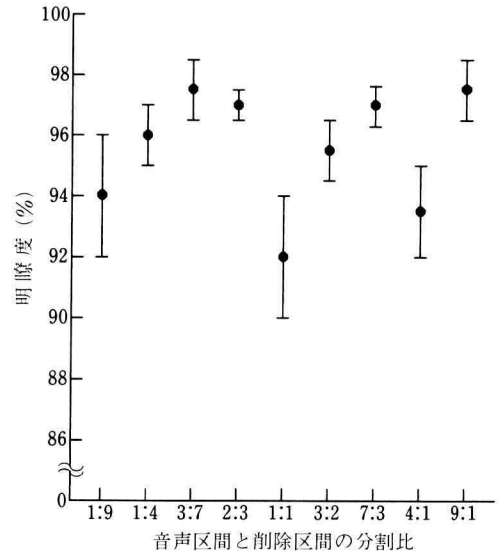


Fig. 4 Intelligibility against Ratios of Speech to Erasing Durations

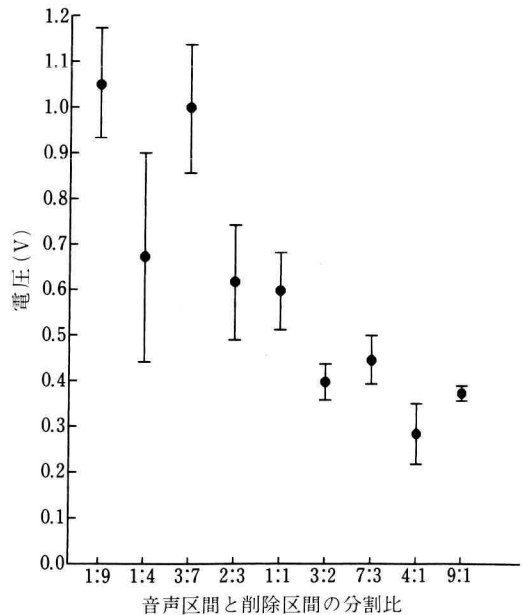


Fig. 5 Relations between Dividing Ratios and Output Voltages of Speech Signal V_{p-p}

として音声波を分割した。分割の割合は、1:9 から 9:1 までの 9 通りについて行った。その結果、明瞭度は 92% から 98.5% までと非常に高く分割比による変化があまり見られない。各個人の明瞭度を標準偏差として計算すると、最大でも $\pm 2\%$ 程度である。この実験から、細かく時間間隔の分割を繰り返した場合には、1:9 と音声波の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がわかった。

また、一周期の削除区間が短い場合について、被験者が削除された音声信号を聞きとるときに最適なヘッドホン出力を調べた。実験結果を Fig. 5 に示す。その結果より、削除区間が大きい 1:9 場合には、ヘッドホン出力を大きくし、反対に、音声区間が大きいとれるときには、出力電圧を低くする必要があることがわかった。このことは、削除区間を細かくして全体の削除区間を大きくとった場合には、明瞭度が十分あることによると考えられる。

4. む す び

本実験データに基づいて、音声波と削除区間を合成して受聴した場合の特徴について検討した。削除区間のある音声波を聞きとったときの明瞭度は、各実験条件においてほぼ直線的に低下すること、ならびに音声波の削除区間が長いほうが明瞭度が低下していることを実験より示した。しかし、一周期に含まれる音声波の削除区間の割合が長い場合、例えば 1:9 のような割合の場合でも削除される区間の短いほうの明瞭度が高いことが実験から明らかになった。また、そのとき音声波のヘッドホン出力を調べた結果によると音声波の削除区間が大きいほど大きな出力を必要としているようである。

謝 辞

本実験の被験者各位に感謝いたします。

参 考 文 献

- 1) 例えば、三浦種敏監修：“新版 聴覚と音声”，電子通信学会編（昭和 57 年）。
- 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) 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) .
- 4) 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).
- 5) 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).
- 6) 例えば、樋渡涓二編著：“視聴覚情報処理概論”，pp. 115-121, 昭晃堂 (1987).
- 7) 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”，第 4 回ヒューマン・インタフェース・シンポジウム論文集, pp. 321-324 (1988).
- 8) 桑原尚夫, 境 久雄：“連続音声の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).
- 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) 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）。