

〈一般研究課題〉 雑音環境における音声の明瞭性改善に関する研究

助成研究者 愛知県立大学 金森 康和



雑音環境における音声の明瞭性改善に関する研究

金森 康和
(愛知県立大学)

Study on improving clarity of speech in noisy environments

Yasukazu Kanamori
(Aichi Prefectural University)

Abstract :

As the conditions of balance of phoneme number, the difference of male and female speakers, and effect of route and SNR, a noise-added speech obtained by adding noise has been created. Twelve men and women was evaluated in the harsh sense of recognition of voice. As a result, that a certain relationship to the average sound pressure level and the recognition rate was found. We found that phoneme with low recognition rate is voiceless plosive such as / t / and / p /. Further, the nasal sound such as / n / and / m / are also with low recognition rate. In addition, a result that male speaker is higher than the female speaker in harsh feeling was obtained. According to our analysis, it is considered that the increase in sound pressure level is one of the factors of discordant. By analysis of the characteristics of the low recognition rate and harsh phonemes, we found that the types of plosive sound is greatly concerned with the results. As the presentation and control method of the desired sound, we have studied the preferred physical characteristics of the horn in everyday life. Good impressions and alarm feelings of a horn becomes a trade-off, the current horn as a sign is evaluated by using the physical characteristics to determine its suitability. It is believed that we can increase the alert feeling by reducing the difference between the first and second spectral peak on the frequency axis and the fundamental frequency.

1. はじめに

人間は、日々の生活環境の中で音情報を聴いて様々な感情変化を起こしている。無意識のうちに音情報を聴いて楽しくなったり興奮することもある。逆に、落ち着いたり悲しくなったり、特殊な雑音でいらいらすることもある。特に、雑音の大きい環境ではアナウンス等の音声情報が伝わらないだけでなく、聴力が著しく低下することもある。また、音楽療法等に用いられることから分かるように、音情報は人の心理面だけでなく、健康面にも影響を与えている。

本研究では、まず、日常生活に直面する音・音声情報・騒音について、聴取実験などを通して、これらの情報に含まれる音響・韻律情報の分析を行い、基本周波数やパワー、異なる周波数帯域の振幅成分の構成情報などが人間の知覚・情動に与える影響の度合い・分布状況を系統的に解析する。次に、雑音環境下において、音声の明瞭性を改善するための方法を検討する。これらの知見を利用して、日常生活において望ましい音の提示・制御法を明らかにする。

2. 地下鉄車内の音声明瞭度の改善

2.1 騒音分析による異なる騒音に適応可能な音声強調手法の検討

バスや鉄道、タクシーなどの公共交通機関は幅広い年代層の人々に多く利用されている。そのため、初めて訪れた場所において公共交通機関を利用する機会も少なくない。その際、駅名や進行方向など移動をする時に必要な情報を提供してくれる音声案内は重要となる。しかし列車の車内では、激しい騒音が発生することで、音声案内が聞き取りづらくなる。本研究では騒音が大きい環境下での音声案内の代表例として、地下鉄の車内アナウンスを取り上げる。

先行研究として、尾林ら(1)の研究より、地下鉄の車内音声进行分析した結果、車内音声の高周波数成分が約3.8kHz部分を境に急激に失われていることが分かった。そこで、音声の可聴性が損なわれている原因は、帯域制限にあると仮定し、可聴性改善のために音声の帯域を広げる手法を提案した。聴覚実験の結果、音声の帯域を広げる手法によって可聴性の改善を図ることが出来た。しかし、この手法では、一部の子音に関して十分な有効性が得られていないという課題点が残った。子音の可聴性が改善されなかった原因を探り、子音の振幅が騒音の振幅に対して大幅に小さいことが原因であると仮説を立てられ、SNRを変えて音声の振幅を強調する手法を提案し、聴覚実験を行った結果、子音の認識率を約10%改善することができた(2)。

これまでに車内アナウンスの可聴性を改善するために提案されてきた手法は、ある一つの区間の騒音に対してのみの結果である。地下鉄の路線は他にも多く存在し、全ての区間の騒音が同じ音声的特徴を示すとは限らない。そのため、提案されてきた手法が地下鉄の路線が変更されても同じように有効な結果をもたらすかどうかは分からない。名古屋市営地下鉄に焦点を当て、各路線の騒音に対して音響分析を通して傾向を探り、提案手法の有効性について検討する必要がある。

2.2 マルチ区間の騒音分析

名古屋市営地下鉄における5つの路線における4～6の計24個の騒音区間に対して分析を行う。提案していた手法はSNRを用いて振幅強調を行うため、騒音と音声の振幅は一定となる。そのため、騒音の周波数に着目し、パワースペクトルを求めることとした。結果の一部をFig.1、Fig.2に示す。Fig.1.は同路線における6区間の騒音パワースペクトル結果である。低周波数成分に主なパワーを持っており、ピーク周波数の傾向が似ていることが分かる。Fig.2.は異なる路線同士の騒音

パワースペクトル結果である。主なピークの周波数帯域が異なっていることが分かる。

パワースペクトル分析から得た結果における各騒音の第1ピーク、第2ピークをパラメータとして、クラスター分析を行い、騒音のグルーピングを行った。結果をFig.3.に示す。Fig.3.より、10種類に分類することができた。このように、24区間の騒音に対してパワースペクトル分析を行った結果、スペクトルの傾向が異なる場合があることが分かった(2)。つぎに、大きく10種類の騒音を用いた音声聞き取りの実験を行い、騒音付加音声による音声強調手法の有効性を確認する。

聴取実験からは最適SNRを求める。最適SNRとは、音質が劣化していない最大のSNRであるが、これらは聴取実験における耳障り感と音量の評価(5段階SD法、耳障り感：感じない、あまり感じない、少し感じる、やや感じる、とても感じる。音量：小さすぎるから非常に大きい5段階)を基に考えられる。この2つの評価値のうち音質評価としては耳障り感の方を重要視する。音量についての評価は3.5以上のとき、耳障り感についての評価は2.5以上を問題あるとする。よって、音量は問題あるが耳障り感について問題がないとするときは、音質劣化とはしない。逆に音量に問題がなくても耳障り感に問題がある場合は、音質劣化であるとする。もしこのときのSNR=-5[dB]であれば、最適SNRは-10[dB]であると定義する。

この最適SNRと認識率の関係は以下のFig.4 ～ Fig.6のようになる。以上の結果から、最適SNRにおける認識率が70%を下回る路線があることがわかる。また、音圧レベルを大きくすることで認識率も向上する傾向にあるが、多くの場合音質が劣化してしまうというトレード・オフの関係でもある。どの音素が耳障りとなっているかを検証するために小規模な聴覚実験を行った結果、子音 /s/, /z/ が耳障りとなっていることと、音声認識率改善の妨げとなっていることがわかった。

2.3 騒音下音声の耳障り感の要因分析

以上のように、名古屋市営地下鉄の主要5路線についても考慮し、SNRを変化させて振幅強調の手法を提案し、有効性を図ってきた。しかし、聴覚実験を行う中で、振幅強調に一つの欠点があることがわかった。振幅強調をする際の音質劣化である。認識率の改善が見られない音素に対しては音圧レベルを上げることで認識率向上につながるが、音質劣化により耳障りと感じることがあることがわかった。この耳障りが発生する原因を探る必要がある。

先行研究と同様に、音声に騒音を付加させた、騒音付加音声を用いた。また、音素数のバランスや、男女話者の違い、SNRや路線による影響などを条件として設定した。作成した騒音付加音声

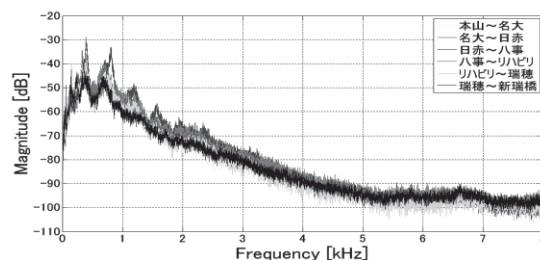


Fig.1. Power spectrum of noise of same line

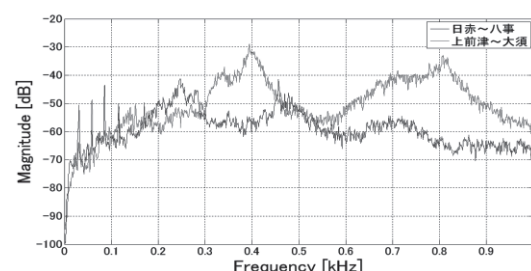


Fig.2. Power spectrum of noise of different line

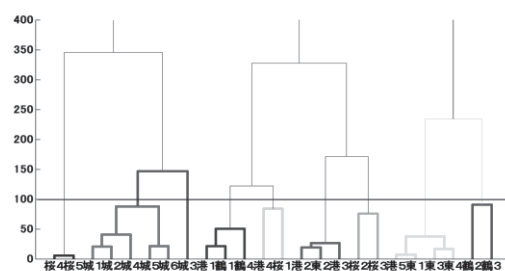


Fig.3. A cluster dendrogram of noises

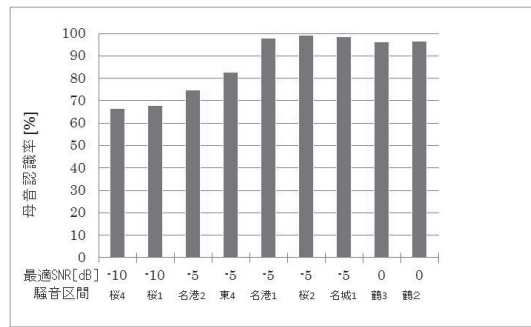


Fig.4 Vowel recognition rate in the optimum SNR

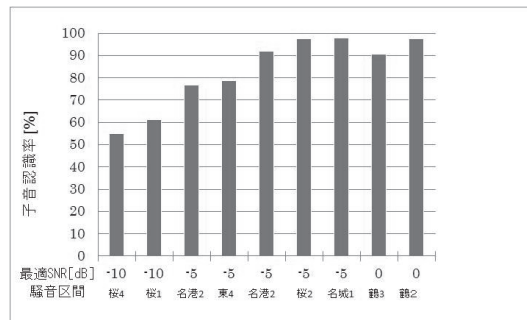


Fig.5 Consonant recognition rate in the optimum SNR

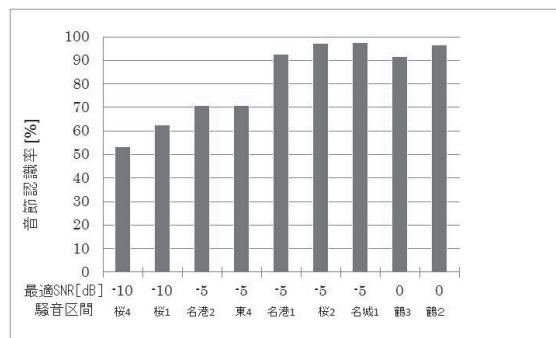


Fig.6 Syllable recognition rate in the optimum SNR

は、男女話者1名ずつ、3段階のSNR、地下鉄騒音5路線に単語30単語を組み合わせた900単語としている。これを話者ひとりあたり300単語になるよう、ランダムに選出している。実験の被験者は健常な聴力を持つ20代の男女8名である。音声の認識と耳障り感や音量の評価を行ってもらった。

3. 実験結果

Fig.7に音素ごとの耳障り感の結果を示す。男性話者の方が耳障りに聞こえていることがわかる。また、音素でみると、/b/や/g/が特に耳障りに聞こえることがわかる。これらは、有声破裂音ということが共通している。SNRと耳障り感の関係をFig.8に示す。SNRが大きくなるにつれて耳障り感も大きくなる傾向があることがわかる。これに対応した、音声認識率については、SNRが大きくなるにつれて音声のレベルが高く、認識率は向上する傾向にあることがわかった。SNRが-10dBの時は、0dBに比べ、86%からおおよそ7%音素認識率が向上する。耳障り感の要因を分析する

中で、耳障りになりやすい音素があることがわかった。それは/b/や/g/であり、ともに有声破裂音であることから、有声破裂音が耳障りになる要因の一つと考えられる。また、SNRが大きいと耳障り感も大きくなる傾向があることがわかり、認識率も向上することがわかった(4)。

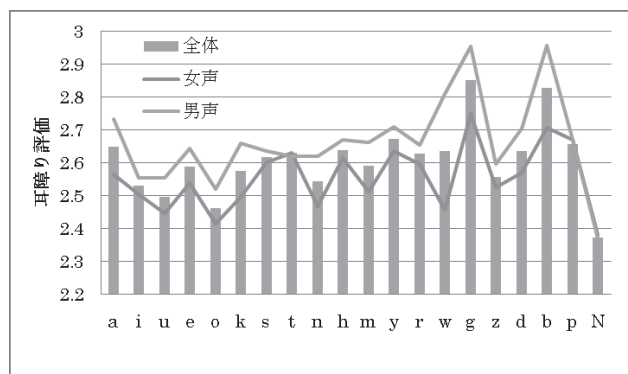


Fig.7 Relations of a phoneme and the harsh evaluation level

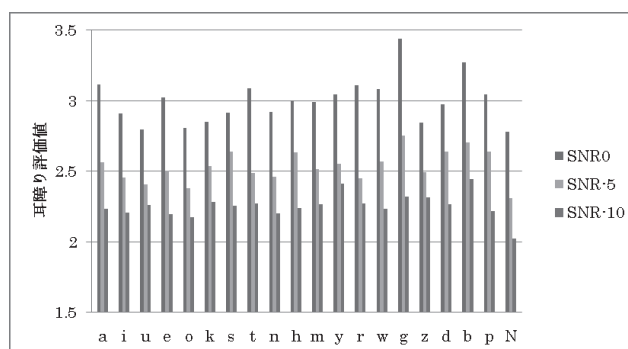


Fig.8 Difference in harsh feeling by SNR

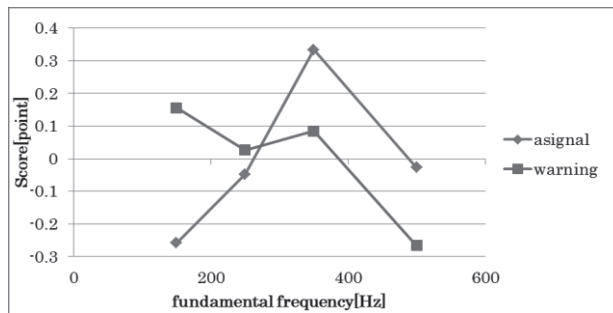
3. スペクトル構造が自動車クラクションの印象に与える影響

日常生活において望ましい音の提示・制御法を明らかにする一例として、自動車クラクションを考える。現在、どの車にもクラクションが搭載されている。高田らの調査(5)では140名にアンケートを行ったところ、現行のクラクションの音量について「大きい」と感じる人が全体の42%,音質については34.8%の人が「耳につき、聞くと不快に感じる」と回答している。クラクションの警報特性を調べる為に我々が行った先行研究(6)では音量99～112dB, 基本周波数250Hz, 500Hzに設定し、周波数帯域幅が大きく、かつ低・全帯域間の音量比が1.1を超えないスペクトル構造をもつクラクションが適切であると確認された。本研究では先行研究(6)で検討されなかった基本周波数、スペクトル第一ピーク(以下P1と統一する)の周波数と音量に注目し、クラクションの音質を改善する条件を求めることを目的とする。

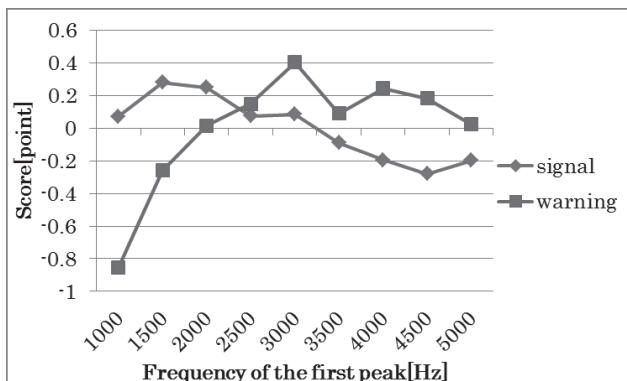
3.1 試験音

先行研究(6)で最も好印象を抱かれたクラクションを原音とし、基本周波数、P1の周波数・音量の情報を変更した。先行研究^[2]で使用された250Hz, 500Hzに新たに150Hzと350Hzを加えて計4種類の基本周波数をTD-PSOLA法を用いて変換した。次にP1の周波数はP1のフォルマント情報を1000Hzから500Hz刻みで5000Hzまで9段階変更する。さらに、P2のパワーを基準にP1のパ

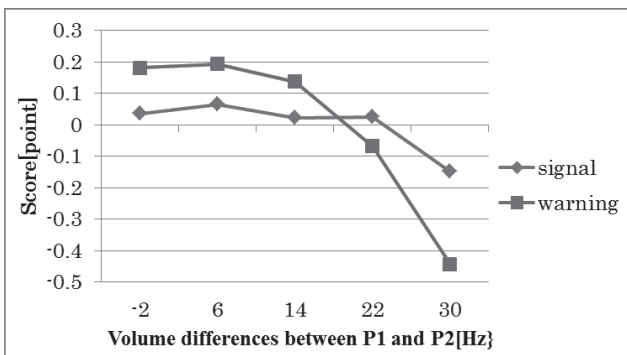
ワーを−2dBから+8dB間隔で30dBまでの5段階変更し、合計180個の試験音を作成した。



(a) Relationship between fundamental frequency and Relevance score



(b) Relationship between frequency of P1 and Relevance score



(c) Volume differences between P1 and P2

Fig.9 Relevance score

10対の形容詞尺度などと「クラクションの警報として適切である－不適切である」, 「クラクションの合図として適切である－不適切である」を用いて、7段階のSD法により聴覚実験を行った。被験者は正常な聴覚を持つ10名(男性7名, 女性3名)であり、簡易防音室でヘッドフォンを着用し、被験者が望めば何度でも試験音を聞ける条件にした。

3.2 結果と考察

「クラクションの警報として適切である－不適切である」, 「クラクションの合図として適切である－不適切である」にランキット変換を行い、被験者の平均評価点を各条件に注目した結果をFig.9に示す。Fig.9に分散分析を行った結果、基本周波数350Hz, P1の周波数1500, 2000Hzの時、合図として特に適切であると判断されることがわかった。但し、P1のパワーを上げすぎると(+30dB)合図として不適切であると判断されることがわかった。基本周波数が350Hz以下, P1が3000Hzの時、警報として適切であると判断されることがわかった。但し、警報としての適切度も

P1のパワーを上げすぎる(+22dB以上)と合図と同様に警報として不適切であると判断されることがわかった。今後更に詳しく分析する。

謝辞

この研究成果は、本研究室学生堀口元気氏、高山雄氏の協力を得ている。ここに記して謝意を表する次第である。

参考文献

- (1) 尾林、金森：帯域制限された地下鉄車内アナウンスの可聴性、日本音響学会春季研究発表 1-3-3,2008
- (2) 石樽、金森：音声強調を用いた地下鉄アナウンスの可聴性の改善について、電子情報通信学会技術研究発表、SP-110 107-112,2010
- (3) 石樽、金森：騒音分析による異なる騒音に適応可能な音声強調手法の検討電気関係学会東海支部連合大会、Q4-5, 2011
- (4) 高山、金森：騒音下音声における耳障り感の要因検討、電子情報通信学会2014年総合大会、D-14-7,2014.3
- (5) 高田、他:自動車警笛音に関する意識調査,環境工学総合シンポジウム講演論文集,No17,pp.103-106, 2007
- (6) Genki Horiguchi, Yasukazu Kanamori:”Evaluation of the car horn based on its physical properties”, SISA2013, RS3-11(2013.10)
- (7) 堀口、金森：スペクトル構造が自動車クラクションの印象に与える影響、電子情報通信学会2014年総合大会、D-14-8,2014.3

