



平成27年度研究助成 【サウンド技術振興部門】より

## 屋外拡声音声品質の文章了解度を ブラインド予測する手法の開発

独立行政法人 国立高等専門学校機構  
都城工業高等専門学校 電気情報工学科

助教

小林 洋介

### 1. 屋外拡声器について

「屋外拡声器」と聴いて図1の装置がパツと思いつくだろうか。この装置は「防災無線」の俗称で地方部では広く利用されており、毎日の時報をこの装置から流している自治体も多い。拡声装置に入力される音声が無線伝送されていることが俗称の由来だが、音声を通信して放送するシステムの末端部にあるのがこの拡声器である。都市部でも駅や空港、学校での各種放送のために、あらゆるところに設置されている。しかし、重要な放送のための装置である反面、拡声器から放送される音声は「何を言っているのかよくわからない」ことが多いという致命的な欠陥がある。

2011年の東日本大震災後の調べでは54%の人



図1 都城高専に設置されている屋外拡声器

が震災時に大津波情報を屋外拡声器からの放送により聴取していたものの、20%の市民がよく聴き取れなかったと回答している。屋外拡声器からの放送により命が助かった人がいた半面、聴き取れば助かった命があったかもしれないと考え、改善していく必要がある。そこで、2013年に屋外拡声器の問題を解決するため、日本音響学会内に災害等非常時屋外拡声システムのあり方に関する技術調査研究委員会が設置され、2015年3月に技術基準案が策定された<sup>1)</sup>。

### 2. 音声の明瞭度・了解度とその予測法

この規準案では、聴取者に伝搬される最終的な拡声音声の品質を、音声の主観評価（聴取試験）による明瞭性評価によって求めることが記載されている。音声の明瞭度評価の歴史は古く、1929年にFletcherらによる単音節明瞭度試験が開発<sup>2)</sup>された。その後、単語を用いた了解度試験として単語の語頭1音素の聴き分けであるDRT（Diagnostic Rhyme Test）が開発され、アメリカでは技術標準となっている<sup>3)</sup>。明瞭性の評価指標は単音節を用いた場合を明瞭度（articulation）、単語または文章を用いた場合を了解度（intelligibility）と呼び、複数の単音節または単語を複数の被験者に提示し、その平均正答率を求める。日本語では、発話する単語の言語的な難易度を統制した親密度別単語了解度試験法エラー！参照元が見つかりません。と

英語版DRTを日本語に置き換えた日本語DRTエラー！参照元が見つかりません。等があり、音声システムやホール設計で使われている。

了解度試験は1音場条件につき少なくとも20単語程度が必要であり、複数の音場条件を評価すると被験者1名あたり数百単語となるため、実験総数は数千単語となり、聴取実験の規模が大きく、評価コストが膨大になる。よって数値解析による了解度の予測が多く、研究者によってなされている。既存の音声明瞭度・了解度を予測する物理的な指標の一つに主として屋内での明瞭度指数であるSTI (Speech Transmission Index) がISOとIECで技術標準<sup>5)</sup>となり、計測器も発売されている。しかし、STIは基準信号を放送し、拡声系の伝送特性より計算際に伝送による信号劣化を全て聴取性能劣化として計算を行うため、基の音声が入りこみと聴き取りにくくても、無歪で伝送できればSTIは最も高いスコアになる。このため、STIには発話速度や発話された信号のパワー調整といった音声信号処理による了解度改善効果は評価できないという問題がある。よって、STIに代わる了解度予測法の研究が盛んであり、音声認識による了解度予測法<sup>7)</sup>などが研究されているが手法は確立されていない。

### 3. 屋外拡声器の了解度とその予測

屋外拡声器の品質に関する先行研究に戸井田

の例エラー！参照元が見つかりません。があるが、研究活動が盛んになったのは東日本大震災以後である。特にロングパスエコーと呼ばれる、エコーのように遅れて聴こえてくる他の拡声器からの妨害に関する聴取実験が多数報告されている<sup>8),10),11)</sup>。筆者らもロングパスエコーを含む伝送品質劣化について主観評価を行い研究報告している<sup>12),13),14)</sup>。さらに筆者は、主観評価による拡声音声の了解度の基準信号を利用しない予測に取り組んでいる。これまでに、IP電話のシングルエンド品質評価 (受信音声のみからの品質評価) に用いられるITU勧告P.563の3SQM (Single Sided Speech Quality Measure)<sup>15)</sup> で利用されている内部特徴量を機械学習することで予測することを試み、ロングパスエコーを考慮した文章による了解度評価の予測精度は低い<sup>12)</sup>ものの、屋外拡声器の伝送特性のみを考慮した単語了解度は高精度に予測することが可能であった<sup>14)</sup>。

機械学習を用いた単語了解度の予測フローを図2に示す。この予測法は、学習条件の被験者ごとの反応を教師とした判別器を構成し、疑似的に被験者実験を再現する。図3にこの手法を用いた単語了解度の予測結果を示す。図の縦軸が主観了解度であり、横軸が予測了解度である。丸プロットが学習に用いた音声信号、三角プロットが学習に用いなかった音声信号の主観評価結果をそれぞれ予測した結果である。図より学習に用いなかった条件も一部を除いて対角線上

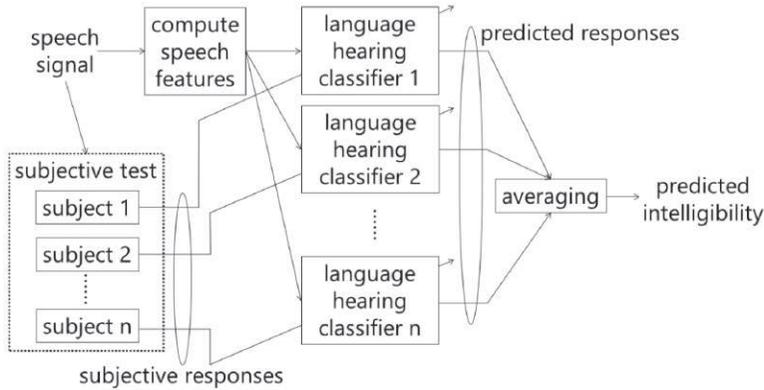


図2 単語理解度の予測フロー

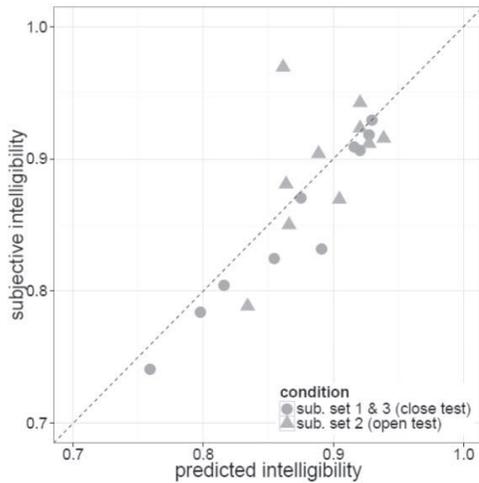


図3 機械学習を用いた単語理解度の予測結果

にあり、高精度な予測ができています。この時の主観評価値と予測値の平均二乗誤差は4.5%であり、2つの主観評価結果間の平均二乗誤差9.3%を下回る制度を達成した。

#### 4. 今後の展望

単語理解度に関しては高精度な予測が可能となったが、ロングパスエコーは単語の音声長では評価できないため、複数の単語または文章音声で聴取実験を行う必要がある。筆者が過去に検討した際に、音響信号から直接文章音声了解

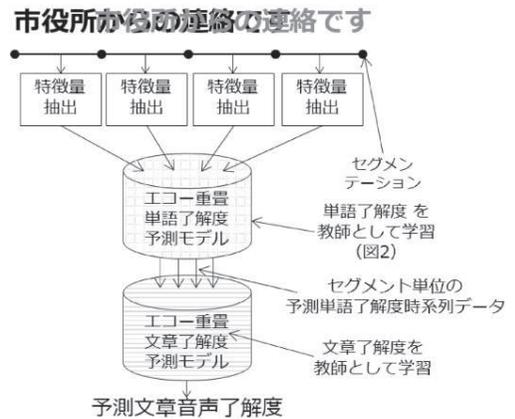


図4 予測単語理解度の時系列データを用いた文章音声了解度予測モデル

度を予測する統計モデルでは主観値と予測値の相関係数が0になるなど非常に精度が低かった<sup>12)</sup>が、図2の単語理解度予測モデルが高精度であったことから、図4のように文章音声をセグメンテーションし、セグメントごとの予測単語理解度値を求め、その時系列データから文章理解度予測モデルを構築する手法を検討する。本稿の執筆時点では、文章理解度予測モデルの構築に利用する文章音声の了解度の主観評価を行っており、主観値が集まり次第、統計モデルを構築する予定である。

## 参考文献

- 1) 災害等非常時屋外拡声システム性能確保のための基準案および解説, <http://asj-disaster-prevention.acoustics.jp/>, 2015
- 2) H. Fletcher and J.C. Steinberg : Articulation Testing Method, Bell Syst. Tech. Journal, 1929
- 3) ANSI/ASA S3.2-2009(R2014) : Method for Measuring the Intelligibility of Speech over Communication Systems, 2009
- 4) 坂本修一, 鈴木陽一, 天野成昭, 小澤賢司, 近藤公久, 曾根敏夫 : 親密度と音韻バランスを考慮した単語理解度試験用リストの構築, 日本音響学会誌, 54(12), pp.842 - 849, 1998
- 5) 近藤和弘, 泉良, 藤森雅也, 加賀類, 中川清司 : 二者択一型日本語音声理解度試験方法の検討, 日本音響学会誌, 63(4), pp.195 - 205, 2007
- 6) ANSI S3.5-1997 (R2012) : Methods for calculation of the speech intelligibility index, 1997
- 7) Y. Takano and K. Kondo : Estimation of Speech Intelligibility Using Speech Recognition Systems, IEICE Trans. Inf. and Syst., vol. E93-D, no. 12, 2010
- 8) 戸井田義徳 : 屋外音場における明瞭度, 日本音響学会誌, 43(7), 519 - 525, 1987
- 9) 虎井駿, 崔正烈, 坂本修一, 佐藤逸人, 森本政之, 鈴木陽一 : 親密度を統制した4連単語によるロングパスエコー下の理解度, 信学技法 EA2013-54, pp.7 - 11, 2013
- 10) 宮下知理, 崔正烈, 坂本修一, 森本政之, 鈴木陽一 : 単語間ポーズがロングパスエコー環境下の単語理解度に及ぼす影響, 信学技法 EA2014-21, pp. 55 - 58, 2014
- 11) 飯田一博, 野村宗弘, 石井要次, 大島俊也, 内藤大輔 : バイノーラル再生による屋外防災放送の単語理解度の再現精度, 日本音響学会聴覚研究会資料, 45(5), pp.413 - 420, 2015
- 12) 小林洋介, 高橋憲敬, 近藤和弘 : ロングパスエコーを想定した文章理解度評価法と伝送パラメータを用いない推定の基礎検討, 信学技法 EA2014-13, pp.7 - 12, 2014
- 13) 西田楓, 小林洋介, 近藤和弘 : 音声伝送系の品質劣化を考慮した屋外拡声器の理解度評価, 音講論(春), pp.543 - 544, 2015
- 14) 小林洋介, 近藤和弘 : 判別器を用いた屋外拡声音声理解度の予測法, 信学技法 EA2015, 2015 (発表予定)
- 15) ITU-T Rec. P.563 : Single-ended method for objective speech quality assessment in narrow-band telephony applications, 2004