



単語了解度によるスピーチプライバシーの評価 : 音の空間特性及び時間特性の影響

星野, 康
森本, 政之
佐藤, 逸人
佐藤, 洋

(Citation)

神戸大学大学院工学研究科・システム情報学研究科紀要, 4:1-6

(Issue Date)

2012

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(URL)

<https://hdl.handle.net/20.500.14094/81004234>



単語了解度によるスピーチプライバシーの評価 —音の空間特性及び時間特性の影響—

星野 康^{1*}・森本 政之²・佐藤 逸人²・佐藤 洋³

¹ 工学研究科建築学専攻/日本板硝子環境アメニティ(株)

² 工学研究科建築学専攻

³ 産業技術総合研究所

(受付: June 11, 2012 受理: August 1, 2012 公開: August 16, 2012)

キーワード: スピーチプライバシー、単語了解度、病院診察室、残響音、空調騒音

筆者らは、先行研究において、遮音性能、暗騒音レベルおよび発話レベルから、単語了解度を予測できる等単語了解度線を提案した(Applied Acoustics, 73, 43-49(2011))。しかし、了解度に影響を及ぼすと考えられる暗騒音の空間特性や、残響音の時間特性および空間特性が考慮されていなかった。本報告では、それらが単語了解度に及ぼす影響を確認し、等単語了解度線の様々な音場への適用性を検討するために、暗騒音の両耳間相関度(ICC)をパラメータとした単語了解度試験と、残響音の ICC および残響音の有無をパラメータとした単語了解度試験の2つの実験を行った。実験の結果、単語了解度は残響音の付加により低下するが、暗騒音および残響音の ICC は影響しないことを明らかにした。この結果より、先行研究で提案した等単語了解度線を、残響音が存在する音場に適用した場合、スピーチプライバシーの観点からは安全側であるが、スピーチプライバシーの考慮が必要な室の設計において、遮音性能あるいは暗騒音レベルについて、最大で10dB程度の過剰設計に繋がる可能性があることを示した。

1. はじめに

近年、我が国では、個人情報保護の需要の高まりもあり、病院の診察室、銀行、薬局、会議室などにおいて、会話による情報の漏洩が問題視されつつある。北米では、スピーチプライバシー[1]あるいはスピーチセキュリティ[2]といった用語でこの問題を表し、古くから研究が進められてきた。Cavanaugh[1]は、スピーチプライバシーの問題には、隣室から漏れてくる会話音による作業妨害と、隣室に会話音が漏れることによる情報漏洩の2つがあると主張している。ここでは、後者の問題について論じる。

スピーチプライバシーの考慮が必要な室の設計において、要求される性能を確保するためには、会話による情報漏洩の程度と設計時に考慮できる物理量の関係を知る必要がある。会話による情報漏洩の指標として、過去のスピーチプライバシーに関する研究では、音声刺激が正しく聴き取れた割合である音声了解度が用いられてきた。ANSI (American National Standards Institute) による用語集[3]においても、スピーチプライバシーは "Technique to render speech unintelligible to casual listeners" と定義されており、音声了解度との関係が強調されている。

Parkら[4]は、室の音響物理量として壁の遮音性能に着

目し、STC (sound transmission class) [5]とRw (weighted sound reduction index) [6]と音声了解度の関係について論じた。しかしながら、音声了解度に大きく影響する室の物理量である暗騒音レベルについて、取り扱っていない。そこで、筆者らによる先行研究[7]では、遮音性能と暗騒音レベルの両者をパラメータとした単語了解度試験を行い、遮音性能、暗騒音レベル、発話レベルの3つのパラメータから単語了解度を予測できる等単語了解度線を提案した。

Fig.1に先行研究で想定した状況を示す。話者AとBが個人情報に関わる会話を行っている。隣室にいる第三者Cには、壁を透過した会話音が聞こえる。通常の室内では、会話音について透過音だけでなく、反射音もCに到達するが、先行研究では反射音は考慮していない。Fig.2に先行研究において提案した等単語了解度線を示す。横軸が室の暗騒音のA特性音圧レベル、縦軸は遮音性能を発話レベルの大小を考慮して補正した値である。遮音性能は重み付き特定場所音圧レベル差 ($D_{P,W}$) [8]を用いた。これは、Fig.1に示したBとCの位置における会話音の音圧レベル差 (D_p) (式(1))を、125Hzから2kHzのオクターブバンド毎に算出し、それを基準曲線と比較して求める単一数值評価量である。

$$D_p = L_{S1} - L_{S2}. \quad (1)$$

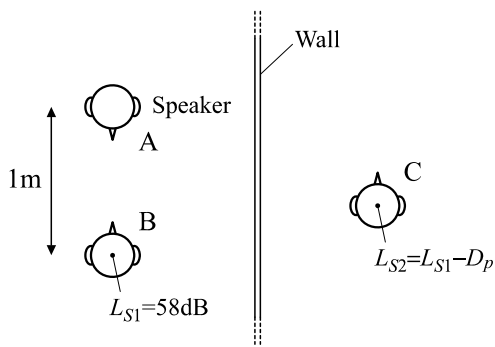


Figure 1 – Situation assumed in the previous study [7]

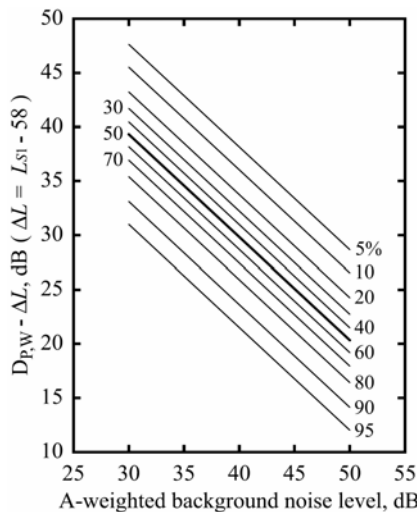


Figure 2 – Equal-intelligibility contours as a function of $D_{P,W}$ and A-weighted background noise level [7].

ここで、 L_{S1} はFig.1に示したBの位置における音声レベル(＝発話レベル)、 L_{S2} はCの位置における音声レベルである。

発話レベルは、Bの位置における会話音のA特性音圧レベルである。発話レベルが高くなると、透過した会話音の音圧レベルも高くなるため、実験で用いられた58dBを基準として、例えば発話レベルが58dBよりも高い場合は遮音性能を低く見積もるよう補正する。この等単語理解度線を用いることにより、設計時に考慮できる物理量から単語理解度を予測することが可能である。

一方、Bradleyら[9]は、音場の空間特性と時間特性がスピーチプライバシーに与える影響について検討し、これらの特性を無視することは室の性能の過剰設計に繋がることを示した。先行研究では、漏れ聞こえる会話音について反射音を考慮していないが、これは音場の空間特性及び時間特性を無視したことと同一である。これは、Bradleyらが主張するように過剰設計ではあるが、スピーチプライバシーの観点からは安全側であることを踏まえてのことであるが、どの程度過剰なのかは定量的に把握しておく必要がある。さらに、暗騒音についても先行研究では全体から広がって聞こえる空調騒音を想定し、両耳間相関度 (ICC) が低い条件のみを用いているが、騒音を付加することによって、スピーチプライバシーを確保するサウンドマスキングシステムの使用を想定すると、

音像がシャープで、特定の方向から聞こえる、ICCが高い条件も考慮する必要がある。

本研究では、先行研究で提案した等単語理解度線が、空間特性や時間特性が異なる様々な音場において、どの程度適用できるかを検討するために2つの実験を行った。実験1では、暗騒音のICCをパラメータとし、暗騒音の空間特性が単語理解度に及ぼす影響について検討した。実験2では、残響音の空間特性と時間特性が単語量理解度に及ぼす影響について検討した。具体的には、残響音のICCと残響音の有無をパラメータとし、単語理解度試験を行った。

2. 実験 I : 暗騒音の両耳間相関度の影響

2. 1 実験方法

実験は無響室で行った。Fig.3にスピーカの配置図を示す。試験用音声として、親密度が高い140単語を用いた。試験用音声は、親密度で統制された単語リスト[10]に含まれる単語のうち、親密度が最も高いものを抽出した。試験用音声の長さは全て4モーラである。女声により無響室録音された試験用音声を音声刺激として用いた。天野ら[10]は、親密度が高いほど単語理解度が高くなることを示した。従って、親密度が高い単語を試験用音声として用いれば、スピーチプライバシーの観点からは安全側の評価が得られる。音声刺激は、被験者の正面に配置したスピーカから提示した。音声刺激のレベルは、被験者の頭部中心に相当する位置で、被験者がいない状態で測定した。A特性音圧レベルの最大値(動特性Slow)を測定し、それぞれの音声刺激について55dBに揃えた。

暗騒音として、空調騒音のモデルとして提案されたHothスペクトル[11]を持つ定常騒音を用いた。暗騒音は3つのスピーカから同時に提示した。なお、暗騒音の両耳間相関度 (ICC_N)を制御するために、それぞれのスピーカから互いに無相関な定常騒音を提示した。正面から提示される暗騒音の音圧レベルを基準として、左右のスピーカから提示される暗騒音の音圧レベルを制御することにより、ICC_Nを制御した。ICC_Nは、ダミーヘッド (KEMAR)を用いて100Hzから10kHzの広帯域のICC_Nを測定し、ICC_Nがおおよそ0.4、0.7、1.0となるように、左右のスピーカから提示される暗騒音の相対レベルを、-3、-8、-∞ dBに設定した。暗騒音のレベルは、試験用音声と同じ位置において、A特性音圧レベルの中央値を測定した。提示レベルは、55、60、65、70、75 dBとした。これをSN比で表すと、それぞれ -20、-15、-10、-5、±0 dBである。

3条件のICC_Nと5条件のSN比を組み合わせた15音場で実験した。被験者は、聴力が正常である20代の学生40名である。被験者には、提示された音声刺激を聴こえた通りにカタカナで書き取らせた。実験は、20名ずつの被験者が参加する2つのグループに分けて行った。一方のグループでは、3条件のICC_Nと、-15 dBと-5 dBの2条件のSN比を組み合わせた6音場について、音声刺激を10単語ずつ提示した。もう一方のグループでは、3条件のICC_Nと、-20 dBと-10 dBの2条件のSN比を組み合わせた6音場について、音声刺激

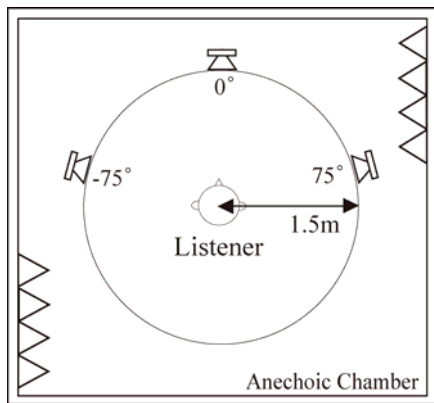


Figure 3 – Loudspeaker arrangement in test I.

を10単語ずつ、さらに3条件の ICC_N と ± 0 dBのSN比を組み合わせた3音場について、4単語ずつ提示した。それぞれのグループにおいて、1人の被験者が同じ単語を聴取することは無く、提示される音場の順番はランダムである。被験者全体で、SN比が-20～-5 dBの音場については200サンプル、 ± 0 dBの音場については80サンプルの回答を得た。

2. 2 実験結果と考察

Fig.4に実験結果を、 ICC_N を関数としてSN比ごとに示す。 ICC_N によらず、SN比が低くなるにつれて単語理解度は大きく低下した。SN比が-20, -15, ± 0 dBの音場では、単語理解度は0%あるいは100%に近い値で一定であり、 ICC_N の影響はみられない。SN比が-10および-5dBの音場では、 ICC_N が0.4と0.7では単語理解度はほぼ等しいが、その2つと比較して、 ICC_N が1.0における単語理解度は低い。SN比と ICC_N を要因とする2要因分散分析の結果、 ICC_N の単純主効果はSN比が-10及び-5 dBで有意 ($p < 0.05$) であった。それらの音場における多重比較の結果、 ICC_N が1.0と他の ICC_N の間のみ有意差 ($p < 0.05$) が見られた。

病院の診察室における暗騒音の両耳間相関度を測定した文献は見当たらないが、筆者らが測定した未発表のデータでは、ある病院の一般的な診察室における ICC_N は、空調騒音のみが音源の場合、0.56から0.74の範囲であった。Fig.4に示した結果より、 ICC_N が0.7以下であれば単語理解度は一定であると考えられる。従って、空調騒音を想定する限りは、 ICC_N は単語理解度に有意に影響しないため、先行研究で提案した等単語理解度線は適用可能であると考えられる。

一方、実験結果より ICC_N が1.0における単語理解度は低下する。ただし、本実験において、試験用音声の ICC も1.0であり、さらに音声と暗騒音が同じ正面方向から到来することに留意する必要がある。この結果はスピーチプライバシーの確保に有効利用できる。例えば、隣室に漏洩した会話音声の ICC が1に近い場合、サウンドマスキングシステムを用いて、音声と同じ方向からマスキングノイズを提示することにより、同じ音圧レベルで比較した場合、空調騒音のように ICC が低い暗騒音よりも、単語理解度を有意に低下させることができるようになる。

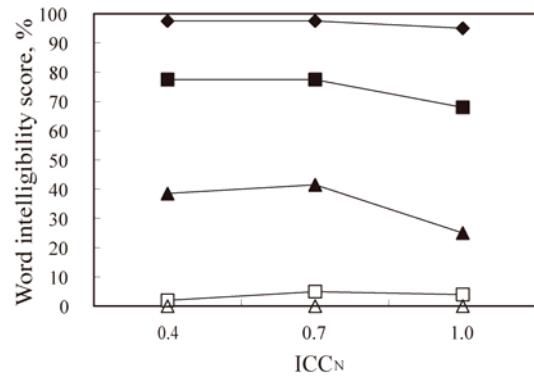


Figure 4 – Word intelligibility scores as a function of ICC_N with a parameter of SN ratio.
(◆:SN ratio = ± 0 dB, ■:-5dB, ▲:-10dB, □:-15dB, △:-20dB)

3. 実験Ⅱ：残響音の両耳間相関度および残響音の有無の影響

3. 1 実験方法

実験は無響室で行った。Fig.5にスピーカの配置図を示す。試験用音声として、親密度が高い120単語を用いた。この120単語は、実験Ⅰで用いた140単語にすべて含まれる。残響音は、残響音の両耳間相関度 (ICC_R) と残響音の付加の有無をパラメータとし、(a) 残響音付加なし、(b) 残響時間1秒の残響音を付加し、 ICC_R を0.44と残響音が全体から広がって聴こえるように設定、(c) 残響時間1秒の残響音を付加し、 ICC_R を1.0に設定、の3条件を用いた。インパルス応答を無響室録音された試験用音声に畳み込むことにより、残響音を付加した音声刺激を作成した。Fig.6に病院の診察室で測定したインパルス応答の例を示す。このインパルス応答は、隣り合った2つの診察室において、一方に音源、もう一方にマイクロホンを設置して測定した。一般に、音楽ホールや講義室などで測定したインパルス応答の場合、直接音が残響音より卓越するが、このインパルス応答では最初に到達する音が壁界によって減衰するため、後続する残響音と比較して弱い。従って、ここでは直接音の無い残響音のみからなるインパルス応答を実験に用いた。

市販のリバーブレータ (YAMAHA SPX900) を用いて作成した残響時間1秒のインパルス応答を残響音の付加に用いた。残響時間の周波数特性及び残響音エネルギーの周波数特性は平坦である。 ICC_R を制御するために、作成したインパルス応答を、残響時間が変わらずかつエネルギーが等分されるように8つに分割した。分割したインパルス応答は互いに無相関である。これらのインパルス応答を試験用音声にそれぞれ畳み込み、1つの音声刺激あたり8個の音源を作成した。条件(b)では、8個の音源を同時にFig.5に示した8個のスピーカからそれぞれ提示した。条件(c)では、8個の音源を同時に正面のスピーカからのみ提示した。条件(b)について、実験Ⅰにおける ICC_N と同じ方法で、 ICC_R をそれぞれの音声刺激について測定した。その結果、平均値が0.44、標準偏差が0.12であった。また、実験者が実際に音

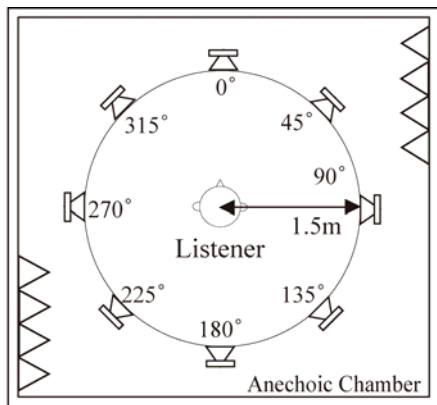


Figure 5 – Loudspeaker arrangement in test II.

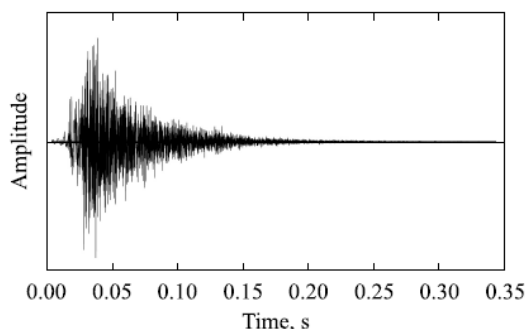


Figure 6 – An example of the impulse response measured in a consulting room of hospital. The sound source was located in the adjacent room.

声刺激を聴取し、全体から広がって聞こえることを確認した。

暗騒音は、実験 I と同じHothスペクトルを持つ定常騒音を用いた。互いに無相関な8個の暗騒音を作成し、同時に8個のスピーカからそれぞれ提示した。ICC_Nは0.53であった。音声刺激と同様に、実験者が定位感が無く全体から広がって聞こえることを確認した。

音声刺激と暗騒音のレベルは、それぞれ実験 1 と同じ方法で測定した。暗騒音レベルは30および50dBとした。音声刺激のレベルは、SN比がそれぞれの暗騒音レベルに対して、-20, -15, -10, -5, ±0dBになるよう設定した。残響音が3条件、暗騒音レベルが2条件、SN比が5条件の組み合わせで30音場を実験に用いた。

被験者は、聴力が正常である20代の学生30名である。被験者には、提示された試験用音声聴こえた通りにカタカナで書き取らせた。30音場に対し音声刺激を4単語ずつ提示した。1人の被験者が同じ単語を聴取することは無く、提示される音場の順番はランダムである。被験者全体で、それぞれの音場に対し120サンプルの回答を得た。

3. 2 実験結果と考察

まず、残響音の両耳間相関度 (ICC_R) の影響について検討する。残響音の条件(b)と(c)の結果を合わせて、Fig. 7にSN比を関数として示す。両者の比較により、ICC_Rが単語理解度に及ぼす影響を検討する。暗騒音レベル50dBの場合

(右図)におけるSN比が0dBおよび-5dBの場合を除いて、ICC_Rが1.0(●)における単語理解度の方が、ICC_Rが0.44(○)の場合より若干高い。それぞれの暗騒音レベルについて、SN比とICC_Rを要因とする2要因分散分析を行った。その結果、暗騒音レベルが30dBの場合(左図)にのみ、ICC_Rの主効果が有意 (p<0.05) であった。しかしながら、ICC_Rが0.44と1.0の単語理解度の差は、最大でも10%程度であり、ICC_Rが単語理解度に及ぼす影響は実用上無視できると考えられる。

次に、残響音が単語理解度に及ぼす影響を検討する。残響音の条件(a)と(c)の結果を合わせて、Fig. 8にSN比を関数として結果を示す。SN比や暗騒音レベルによらず、残響音のない方が、残響時間1秒の残響音がある場合よりも明らかに単語理解度が高い。

先行研究で提案した等単語理解度線は、残響音なしで求められたもので、例えば、設計目標を単語理解度20%に設定した場合、SN比は残響音がない場合は約-17dB、残響時間が1秒の場合は約-10dB となり、残響時間が1秒の場合には約7dB の過剰設計となる。この過剰設計の程度は、設計目標とする単語理解度の値によりやや異なるが、概ね10 dB程度である。スピーチプライバシーを考慮する必要がある空間では、1秒よりも残響時間が長くなることは稀であると考えられる事から、先行研究で提案した等単語理解度線は、遮音性能あるいは暗騒音レベルについて、最大で10 dB程度の過剰設計になる可能性がある。この過剰設計は、スピーチプライバシーの観点からは安全側であるが、効率よく設計するために先行研究で提案した等単語理解度線の補正を今後行う必要がある。先述のとおり、ICC_Rの影響は、ほとんど無視できることから、補正にあたっては残響時間に着目することが重要である。

なお、Fig. 8に示すように、SN比が-20 dBの音場では、残響音の有無やICC_Rによらず、単語理解度は10%以下になる。

4. 結論

本研究では、先行研究で提案した等単語理解度線の適用範囲について検討するために、暗騒音の両耳間相関度と、残響音の両耳間相関度及び残響時間をパラメータとした音場について、単語理解度試験を行った。その結果、以下を明らかにした。

(1) 暗騒音の両耳間相関度の影響は、SN比によって異なる。SN比が0, -15, -20dBの音場においては、両耳間相関度は影響しない。しかし、SN比が-5, -10dBの音場において、暗騒音の両耳間相関度が1.0の場合、0.7以下の条件と比較して、単語理解度が有意に低下する。一方、両耳間相関度が0.7以下では、単語理解度に有意な差はみられない。

このことから、空調騒音のように両耳間相関度の低い音源のみが暗騒音となる音場には、先行研究で提案した等単語理解度線を適用できるとみなせる。

(2) 残響音の両耳間相関度の影響は、SN比によらず単語理解度に影響せず、実用上無視できると考えられる。

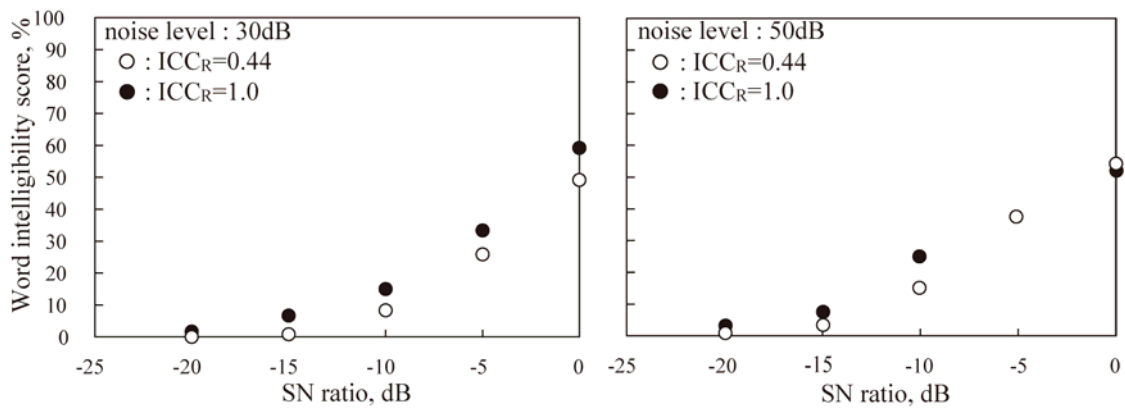


Figure 7- Word intelligibility scores as a function of SNR with a parameter of ICC_R
(Left panel : noise level=30dB, Right panel : noise level=50dB)

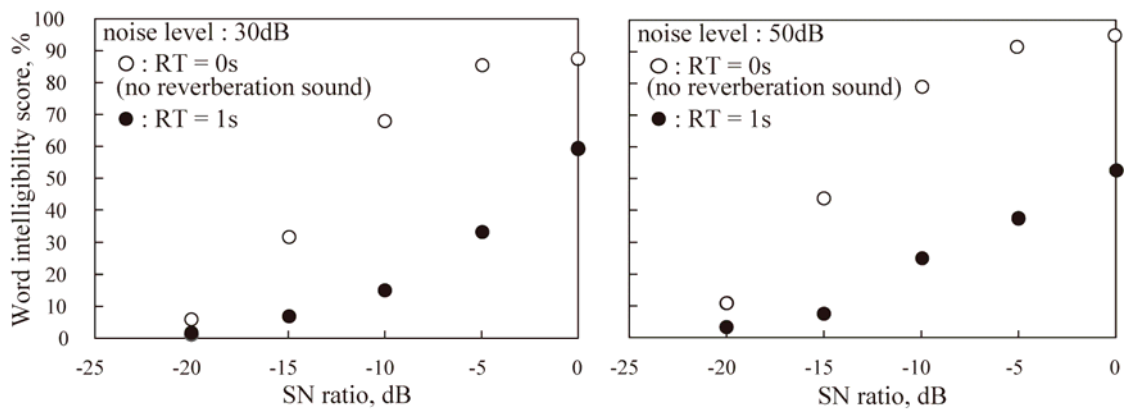


Figure 8 - Word intelligibility scores as a function of SNR with a parameter of RT
(Left panel : RT=1s, Right panel : RT=0s)

一方、残響時間1秒の残響音を付加した音場では、残響音を付加しない音場と比較して、SN比が同じでも単語理解度は有意に低下する。以上のことから、先行研究で提案した等単語理解度線を残響音が存在する音場に適用した場合、スピーチプライバシーの観点からは安全側であるが、遮音性能あるいは暗騒音レベルは最大で10dB程度の過剰設計に繋がる可能性がある。

(3) SN比が-20 dBの音場では、残響音の有無や ICC_R によらず、単語理解度は10%以下になる。

[謝辞] この研究の一部は、科学研究費補助金基盤研究B(20360262)の助成を受けて行った。また、この実験を実施するにあたり、神戸大学工学部の学生であった大谷宗市氏、岡崎啓佑氏に多大なご協力を頂いた。ここに記して感謝の意を表す。

Literature Cited

[1] W. J. Cavanaugh, W. R. Farrell, P. W. Hirtle, and B. G. Watters, "Speech privacy in buildings," *J. Acoust. Soc. Am.*, 34, 475-492 (1962).
[2] B. N. Gover and J. S. Bradley, "Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms," *J. Acoust. Soc. Am.*, 116, 3480-3490 (2004).

[3] ANSI T1.523-2001, *Telecom Glossary*, 2001.
[4] H. K. Park, J. S. Bradley, and B. N. Gover, "Evaluating airborne sound insulation in terms of speech intelligibility," *J. Acoust. Soc. Am.*, 123, 1458-1471 (2008).
[5] ASTM E413-10, *Classification for Rating Sound Insulation*, 2010.
[6] ISO 717-1:1996, *Acoustics - Rating of sound insulation in buildings and of building elements - Part 1: Airborne sound insulation*, 1996.
[7] Ha. Sato, M. Morimoto, Y. Hoshino, and Y. Odagawa, "Relationship between sound insulation performance of walls and word intelligibility scores," *Appl. Acoust.* 73, 43-49 (2012).
[8] JIS A 1419-1:2000, *建築物及び建築部材の遮音性能の評価方法 - 第1部: 空気音遮断性能*, 2000.
[9] J. S. Bradley, M. Apfel, and B. N. Gover, "Some spatial and temporal effects on the speech privacy of meeting rooms," *J. Acoust. Soc. Am.* 125 (2009) 3038-3051.
[10] S. Amano, S. Sakamoto, T. Kondo, and Y. Suzuki, "Development of familiarity-controlled word lists 2003 (FW03) to assess spoken-word intelligibility in Japanese," *Speech Commun.* 51, 76-82 (2009).
[11] D. H. Hoth, "Room noise spectra at subscribers' telephone locations," *J. Acoust. Soc. Am.* 12, 499-504 (1941).

The evaluation of speech privacy by word intelligibility -The effect of spatial and temporal aspects of sound fields-

Yasushi HOSHINO¹, Masayuki MORIMOTO², Hayato SATO²,
Hiroshi SATO³

¹ *Graduate School of Engineering, Department of Architecture / Nippon Sheet Glass Environment
Amenity Co., Ltd.*

² *Graduate School of Engineering, Department of Architecture*

³ *National Institute of Advanced Industrial Science and Technology*

Key words: Speech privacy, Word intelligibility test, Consulting room of hospital, Reverberation sound,
Background noise of air conditioning

In the previous study [Applied Acoustics, 73, 43-49(2011)], the authors presented the equal-intelligibility contours that make it possible to predict the word intelligibility score from sound insulation performance, background noise level and speech level. However, the contours do not consider spatial characteristics of background noise and spatial and temporal ones of reverberation sound, which would affect intelligibility scores. In the present study, two intelligibility tests, that composed of interaural cross-correlations (ICC) of background noise, and ICC and reverberation time of reverberation sound as the test parameters, were performed in order to discuss the application of the contours to various sound fields. The results demonstrated that the ICC of background noise and reverberation sound did not affect the scores, while the scores significantly decrease with increasing reverberation time. The result of the present study demonstrated that the equal-intelligibility contours presented in the previous study provide design goals on the safe-side in terms of speech privacy, and either sound insulation performance or background noise level of rooms where speech privacy should be considered may brought to over-design by 10dB at maximum in the case of being applied in the reverberation sound fields.