

一側性難聴者の残響下での文章理解度・方向性マスキング解除の検討 —雑音下・競合音声下聴取の比較—*

☆辻慎也, 荒井隆行 (上智大)

1 はじめに

室内で音声を聴取するとき、我々の耳には音源から直接到来する音だけでなく、部屋の壁などの面で反射した成分である残響が加わって音が到達している。残響時間が長い場合、音声の時間構造が崩れて (self-masking), 先行する音声の残響により後続の音声がマスキングされ (overlap-masking) 音声明瞭度が影響される[1]。残響は特に聴覚障害者にとって、音声明瞭度を低下させる要因となる[2]。

片側耳は正常であるが、対耳側に軽度以上の難聴を有する状態を一側性難聴 (UHL; unilateral hearing loss) という。患耳が重度難聴を呈するとき、片側聾 (SSD; single-sided deafness) として区別することもある。一側性難聴により聴取に困難を要する場面は難聴側聴取・騒音下聴取・音源定位の3点である[3]。

騒音下での聴取について、音声とマスクがそれぞれ別の位置から到来するとき音声聴取の閾値が改善する (方向性マスキング解除: spatial release from masking; 以下 SRM)。SRM の効果を得るためには両耳聴によるキューが重要である[4]。片耳での聴取では、雑音に対し音声が聞こえる側にあるとき信号対雑音比 (以下 SN 比) が改善し、SRM の効果が得られる[5]。一方で音声がその反対側にあるときは、頭部陰影効果が裏目に作用し SRM は負の値を取る[6]。

残響下において、一側性難聴者は雑音下の音声聴取のために両耳聴と比べ、より大きな SN 比を必要とした[7]。雑音下に対して競合音声下の聴取では、マスクがターゲット音声の知覚に違った影響をもたらす[8]。しかし、残響下において、一側性難聴者を対象とした研究は少なく、検討を進める必要がある。

そこで本研究では、残響下において、一側性難聴者の雑音下 / 競合音声下の音声知覚とそれぞれの SRM を検討することを目的と

して、一側性難聴者と両側聴力正常者のモノラル受聴・両耳聴を対象に文章理解度と SRM を測定している。本稿では原稿作成時点までに集まった結果について議論する。

2 方法

一側性難聴者 6 名 (男性 2 名, 女性 4 名; 22-38 歳) が本研究に参加した。本稿では両側聴力正常者のデータのうち、10 名 (男性 5 名, 女性 5 名; 18-24 歳) を対象とした。

対象のうち 1 名の一側性難聴者について、患耳側の純音聴力検査が他の一側性難聴のある参加者と違う傾向を示した (125, 250 Hz で 30 dB HL, 500 Hz で 60 dB HL, その他はスケールアウト) ので、本稿では分析の対象から外した。

分析の対象とした一側性難聴者について、難聴の期間は 20-38 年で、先天的発症が 4 名、後天的発症が 1 名であった。難聴側は左が 1 名, 右が 4 名。良聴耳の平均聴力レベル (四分法) は 20 dB HL 以下で、難聴の程度は高度難聴が 1 名, スケールアウトが 4 名であった。

2.1 インパルス応答の収録

本研究では上智大学 10 号館講堂の両耳インパルス応答[7]の畳み込みにより刺激音に残響を付与した。収録には 12 面体スピーカ (Brüel & Kjaer, OmniPower 4292) からの swept-sine 信号音源と、3 脚にセットしたダミーヘッド (Neumann, KU 81) を用いた。ダミーヘッドに対する音源の角度を変え、複数収録されたインパルス応答のうち正面 (0°)・左右 (±35°) 3 点の両耳インパルス応答を用いた。音源と受聴点の高さは床から約 1.35 m, 距離は約 3.00 m であった。インパルス応答の残響時間は約 1.6 秒であった。

2.2 文章理解度テスト

文章理解度 (speech reception threshold; 以下 SRT) テストでは、ターゲット文章は NTT-AT 音素バランス 1000 文[9]から 1) 話者が男性ア

* An investigation of speech recognition and spatial release from masking under reverberation for those with unilateral hearing loss: A comparison of speech intelligibility under noise and that under competitive speech, by TSUJI, Shinya and ARAI, Takayuki (Sophia University).

ナウンサーで、2) 話速が 7.5–9.0 モーラ/秒、3) 4–5 文節の、3 つの条件を満たす 250 文を選択し、20 文のリストを 12 個、10 文のリストを 1 個作成した。各文では、文節ごとにキーワードを設定した。

雑音下の SRT の測定では、選択した 250 文の長時間平均スペクトルと白色雑音を Praat によりマッチさせた speech-shaped noise をマスクに用いた。競合音声下 SRT の測定では、日本語単一話者オーディオブック・紙芝居朗読音声コーパス (K-KAC) [10] より、男性声優による「赤いろうそくと人魚」「よだかの星」2 つの朗読音声を選択した。各朗読音声について、300 ms をこえる無音の部分は 100 ms 以下になるように波形を編集した。

Matlab により、両耳インパルス応答を畳み込み、ターゲット音声とマスクに残響を付与した。250 文のターゲット音声と競合音声は、振幅を RMS 正規化した。ターゲット音声の立ち上がり立ち下がりには 10 ms のコサイン関数によって時間包絡を付与した。

One-up / one-down の階段法により、50% 閾値 [11] で SRT を測定した。実験参加者には呈示した刺激音声の復唱を求め、各文節に設定したキーワードを 3 つ以上復唱できた場合は SN 比を 2 dB 下げ、復唱できなければ SN 比を 2 dB 上げた。テストの冒頭ではマスクの等価騒音レベルは 65 dB (A)、音声は 75 dB (A) とした (SN 比 +10 dB からスタートした)。雑音下 SRT の測定ではターゲット音声の最大音圧を 75 dB (A)、競合音声下 SRT の測定では最大音圧を 81 dB (A) とした。設定した最大音圧まではターゲット音声を音圧操作して、最大音圧の時はマスクの音圧を操作して SN 比を増減させた。SN 比の増減が 2 回折り返した後、6 回折り返すまで試行を行い、最後の 6 回の SN 比を平均することで SRT を求めた。テスト前に、10 文章リストの音声を使用して練習を行った。

雑音下・競合音声下それぞれで、SRT は SRM をみるために 3 回測定した。1 回目はターゲット音声とマスクの両方に同位置 (正面, 0°)、2 回目はターゲット音声に左側 (-35°)、マスクに正面 (0°)、3 回目はターゲット音声に右側 (35°)、マスクに正面 (0°) の両耳インパルス応答を畳み込んだ。雑音下の測定では 12 の 20 文リストのうち 3 つを、競合音声下の測定

ではそれと重複しない 3 つの 20 文リストをターゲット音声として用いた。

2.3 手続き

実験は上智大学の防音室で行われた。本稿に関わる手続きは 1) 紙面によるデモグラフィック要因の収集、2) オーディオメータ (RION, AA-79S) による純音聴力測定、3) 文章了解度テストであった (同時に gap 課題を行った)。

刺激音の呈示と回答の正誤・刺激音の SN 比の記録には Matlab の UI を用いた。刺激音はコンピュータからオーディオインタフェース (Roland, Rubix24) を介し、ヘッドホン (SENNHEISER, HDA300) によって呈示した。両側聴力正常者のうち半数にはヘッドホンにより両側呈示 (binaural normal hearing; 以下 BNH, $n = 5$) し、もう半数には片側のみ呈示 (monaural normal hearing; 以下 MNH, $n = 5$) した。MNH の呈示側は左右同数・ランダムに選択した。一側性難聴者 (以下 UHL, $n = 5$) にはヘッドホンにより両側呈示した。

3 結果

Fig. 1 に SRT の raincloud plot を記す。統計分析には R (4.2.1) を用いた。SRT を応答変数、聴取条件・その他のデモグラフィック要因を説明変数として、線形混合モデルによるパラメータ推定を行い、変量効果は実験参加者とした。各説明変数の p 値の算出には lmerTest パッケージを用いた。多重比較では、Bonferroni 法で p 値を補正した。各群の SRT の推定周辺平均と 95% 信頼区画 (以下 M , CI) の算出には emmeans パッケージを用いた。Table 1 に各群の SRT の $M \cdot CI \cdot SRM$ を記す。

雑音下での SRT は、雑音とターゲット音声と同位置の時 BNH と MNH ($M = -0.40$ vs. 4.93; $\beta = -5.33$, $s.e. = 1.28$, $t = -4.17$, $p < 0.01$) の間に有意差がみられた。雑音に対して、音声を BNH の左側・MNH の呈示側・UHL の良聴耳側に置いたとき (以下、同側条件)、BNH と MNH ($M = -4.33$ vs. 3.07; $\beta = -7.40$, $s.e. = 1.51$, $t = -4.89$, $p < 0.01$)、MNH と UHL ($M = 3.07$ vs. -2.00; $\beta = 5.07$, $s.e. = 1.51$, $t = 3.35$, $p < 0.05$) の間に有意差がみられた。音声を BNH の右側・MNH の非呈示側・UHL の患耳側に置いたとき (以下、対側条件)、全ての条件間で有意差がみられた (BNH vs. MNH: $M = -4.40$ vs. 11.20; $\beta = -15.60$, $s.e. = 2.21$, $t = -7.05$ $p <$

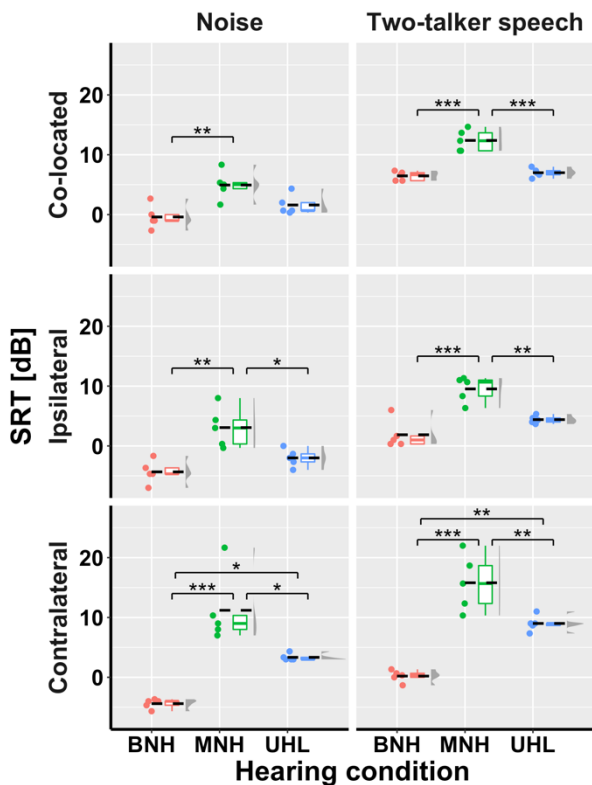


Fig. 1 Raincloud plots for the speech reception thresholds (SRT) including raw jittered data, box-whisker plots, and split-half violins for three hearing conditions; BNH ($n = 5$), MNH ($n = 5$), and UHL ($n = 5$). The solid and broken horizontal lines indicate the median and mean, respectively.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

0.001; BNH vs. UHL: $M = -4.40$ vs. 3.33 ; $\beta = -7.73$, $s.e. = 2.21$, $t = -3.50$, $p < 0.05$; MNH vs. UHL: $M = 11.20$ vs. 3.33 ; $\beta = 7.87$, $s.e. = 2.21$, $t = 3.56$, $p < 0.05$).

競合音声下での SRT は同位置の時 BNH と MNH ($M = 6.47$ vs. 12.40 ; $\beta = -5.93$, $s.e. = 0.76$, $t = -7.45$, $p < 0.001$), MNH と UHL ($M = 12.40$ vs. 7.00 ; $\beta = 5.40$, $s.e. = 0.76$, $t = 7.10$, $p < 0.001$) の間に有意差がみられた。同側条件では, BNH と MNH ($M = 1.87$ vs. 9.53 ; $\beta = -7.67$, $s.e. = 1.19$, $t = -6.41$, $p < 0.001$), MNH と UHL ($M = 9.53$ vs. 4.40 ; $\beta = 5.13$, $s.e. = 1.19$, $t = 4.31$, $p < 0.01$) の間に有意差がみられた。対側条件では, 全ての条件間で有意差がみられた (BNH vs. MNH: $M = 0.20$ vs. 15.80 ; $\beta = -15.60$, $s.e. = 1.82$, $t = -8.79$, $p < 0.001$; BNH vs. UHL: $M = 0.20$ vs. 9.00 ; $\beta = -8.80$, $s.e. = 1.82$, $t = -4.84$, $p < 0.01$; MNH vs. UHL: $M = 15.80$ vs. 9.00 ; $\beta = 6.80$, $s.e. = 1.82$, $t = 3.74$, $p < 0.01$)。

SRM は同位置の SRT から同側 / 対側条件の SRT の差分により求めた (Table 1)。BNH・

Table 1 Estimate marginal means and confidence intervals of the speech reception threshold (SRT) and spatial release from masking (SRM)

聴取条件	マスク (正面)	音声の位置	SRT (M) [dB]	CI		SRM (M) [dB]
				Low	High	
BNH	雑音	正面	-0.40	-2.37	1.57	
		左側	-4.33	-6.67	-2.00	3.93
		右側	-4.40	-7.81	-0.99	4.00
	音声	正面	6.47	5.24	7.69	
		左側	1.87	0.02	3.71	4.60
		右側	0.20	-2.54	2.94	6.27
MNH	雑音	正面	4.93	2.97	6.90	
		呈示側	3.07	0.73	5.41	1.86
		非呈示側	11.20	7.75	14.65	-6.27
	音声	正面	12.40	11.00	13.83	
		呈示側	9.53	7.71	11.36	2.87
		非呈示側	15.80	13.06	18.54	-3.40
UHL	雑音	正面	1.60	-0.37	3.57	
		良聴耳側	-2.00	-4.34	0.34	3.60
		患耳側	3.33	-0.12	6.78	-1.73
	音声	正面	7.00	5.83	8.17	
		良聴耳側	4.40	2.56	6.24	2.60
		患耳側	9.00	6.20	11.80	-2.00

M : 推定周辺平均, CI : 95%信頼区画

MNH・UHL 各群で SRT を応答変数, ターゲットの呈示側・マスクの種類を説明変数としたときの線形混合モデルによるパラメータ推定の結果を Table 2 に示す。各パラメータが負の方向に寄与するときは SRT の低下 (音声明瞭度は上昇する方向に作用される) を, 正の方向に寄与するときは SRT の上昇 (音声明瞭度は低下する方向に作用される) を意味する。

Table 2 Results from regression models analyzing speech reception thresholds under three hearing conditions: BNH ($n = 5$), MNH ($n = 5$), and UHL ($n = 5$)

		推定値	標準誤差	T 値	P 値
BNH	切片	0.09	0.61	0.15	0.89
	位置 (左側)	-4.27	0.48	-9.00	<0.0001
	位置 (右側)	-5.03	0.48	-10.80	<0.0001
MNH	マスク (音声)	5.89	0.39	15.18	<0.0001
	切片	5.58	1.11	5.03	<0.01
	位置 (呈示側)	-2.37	0.66	-3.60	<0.001
	位置 (非呈示側)	4.83	0.66	7.36	<0.0001
UHL	マスク (音声)	6.18	0.54	11.52	<0.0001
	切片	1.39	0.38	3.64	<0.01
	位置 (良聴耳側)	-3.10	0.44	-7.13	<0.0001
	位置 (患耳側)	1.87	0.44	4.29	<0.001
	マスク (音声)	5.82	0.36	16.39	<0.0001

4 考察

本調査では一側性難聴者と両側聴力正常者のモノラル受聴・両耳聴を対象に、残響下において、雑音下・競合音声下の文章了解度テストを行って、一側性難聴者の音声聴取に残響が及ぼす影響、SRMの効果を検討した。

その結果、モノラル受聴 (MNH・UHL) では残響下の音声聴取のため、BNHより大きい音圧を必要とした (Fig. 1, Table 1)。特に聴覚障害者では、残響は音声明瞭度を下げる要因となる[2]。本研究では、一側性難聴においても同様、残響がSRTを上げる要因となることが示された。

線形混合モデルによるパラメータ推定の結果、全ての聴取条件でSRMの効果は有意であったが ($p < 0.001$, Table 2)、モノラル受聴では同側条件のときBNHと比べてSRMの効果が高く、対側条件のときSRMは負の値を取った (Tables 1, 2)。また、マスクが競合音声のとき、全ての聴取条件で雑音の時と比べ有意なSRTの上昇がみられた ($p < 0.0001$, Table 2)。

一方、同じくモノラル受聴であったMNHと比べて、UHLでは競合音声下の同位置条件で有意なSRTの改善を示した ($p < 0.001$; Fig. 1, Table 1)。UHLでは残響下の informational maskingからの解除がみられた。

また、MNHと比べUHLは同側条件でよりSRTが改善し (MNH vs. UHL: -2.37 vs. -3.10)、対側条件でSRTの上昇が少ない (MNH vs. UHL: 4.83 vs. 1.87) と推定された (Table 2)。同側条件・対側条件ともにMNHとUHLの間でSRTには有意差がみられた (Fig. 1)。UHLでは同側条件のみならず対側条件であっても、残響下でのSRMが改善していた。

一側性難聴では両耳間情報と代替し、頭部伝達関数 (head-related transfer function; 以下HRTF) によるスペクトラルキューの活用によって音源方向を判断し得る[12]。過去の検討では一側性難聴者のSRMには難聴期間と有意な正の相関 ($p < 0.05$, $r = 0.37$) がみられ、モノラル受聴でのキューへの適応により残響下のSRMは改善し得ると示唆された[7]。本研究では、両耳聴と違うモノラル受聴におかれたMNHに対して、UHLは28.80年 (ISD: 6.98) と長い期間、一側性難聴の状態だった。MNHとの比較から、発症から長期間経過した一側性難聴者はモノラル受聴でのキューへ

の適応によって、残響下での informational masking から解除され、SRMの効果は改善し得ると示唆された。

本研究の実施当初、ターゲット音声の音圧操作のみでSN比を増減していた。従って、最大のSN比が設定されており (2.2節参照)、いくつかの一側性難聴者の結果はフロア効果でバイアスされた恐れがある。今後、マスクの音圧操作も含めた手続きによりバイアスのないデータを集める必要がある。また、本研究ではダミーヘッドで収録した両耳インパルス応答 (標準化 HRTF) の畳み込みで残響を付与したことにより、MNHとUHLの比較を実現した。しかし、標準化 HRTF の畳み込みで作成した刺激は、本人の HRTF との誤差により実際の聴取とは違ったキューをもつ[13]。今後、実際の音場で残響下の文章了解度テストやSRMの効果を検討する必要がある。

参考文献

- [1] R. H. Bolt and A. D. MacDonald, *J Acoust Soc Am*, 21(6), 577–580, 1949.
- [2] A. K. Nábělek and D. Mason, *J Speech, Language, and Hear Res*, 24(3), 375–383, 1981.
- [3] E. Harford and J. Barry, *J Speech Hear Disord*, 30(2), 121–138, 1965.
- [4] R. Y. Litovsky, *Acoust Today*, 8(2), 18–25, 2012.
- [5] A. W. Bronkhorst and R. Plomp, *J Acoust Soc Am*, 83(4), 1508–1516, 1988.
- [6] A. M. Rothpletz *et al.*, *J Speech, Lang Hear Res*, 55(2), 511–531, 2012.
- [7] S. Tsuji and T. Arai, *Acoust Sci Technol*, (in press).
- [8] N. Corbin *et al.*, *J Speech, Lang Hear Res*, 64(11), 4495–4512, 2021.
- [9] NTTアドバンステクノロジー株式会社, “音素バランス1000文,” 1999.
- [10] 高道他, 情報処理学会研究報告, MUS-131(14), 1–4, 2021.
- [11] H. Levitt, *J Acoust Soc Am*, 49(2B), 467–477, 1971.
- [12] J. B. Firszt *et al.*, *Hear Res*, 319, 48–55, 2015.
- [13] 飯田, 小早川, 日本音響学会聴覚研究会資料, 51(1), 19–25, 2021.