

# 音声波形の時間微細構造から見たアクリル板設置による聞き分け 低下の要因とその簡易改善法

山本 亜優\*<sup>1</sup> 八坂 由衣\*<sup>1</sup> 堀越 哲郎\*<sup>2</sup>

## Undesirable Effect of Acrylic Partitions on Speech Recognition and a Simple Improvement Method Estimated by the Temporal Fine Structure of Speech Sound Waveforms

by

Ayu YAMAMOTO \*<sup>1</sup>, Yui YASAKA \*<sup>1</sup> and Tetsuro HORIKOSHI \*<sup>2</sup>

(Received on Mar. 29, 2023 and accepted on May 25, 2023)

### Abstract

In this study, the phenomenon of acrylic partitions affecting speech waveforms and making them difficult to hear was evaluated by observing speech waveforms from the perspective of the temporal fine structure. The results revealed a strong tendency for the waveform features of voiceless and voiced burst consonants and affricative consonants to be obscured by an acrylic partition. In order to reduce this negative effect on speech recognition, we tested whether the waveforms of these consonants could be restored by forming a small window in the acrylic partition and attaching a splash-proof PVDC film to it. As a result, the waveforms of these consonants were partially but effectively restored. The methods used in this study are simple and inexpensive, yet effective.

**Keywords:** Speech sound waveform, Consonant recognition, Temporal fine structure, Acrylic partition

### 1. まえがき

2019年より世界にパンデミックを引き起こした新型コロナウイルス感染症(COVID-19)は、日本において2023年3月現在も完全には収束していない。その感染対策のひとつとして、対面での会話の際にアクリル板のパーティションを対話者の間に置き、お互いの飛沫を直接浴びないように努める場面も多い。アクリル板はガラス板と同程度の透明度を持ち視覚情報を遮らないこと、比較して軽いために設置しやすいこと、割れた場合での安全性も比較的高いことなどから対面会話時のパーティションとして多く導入されている。一方、アクリル板があることによって音声聞き取りにくくなることも報告されており<sup>1)</sup>、それを補うためにマイクとスピーカーをアクリル板両側に設置する装置も市販されている。このような付加装置はアクリル板の設置に加えての費用や電源の確保が必要となり、教育現場のような場で多数導入することは難しい。そこで本研究では、アクリルパーティションを介した音声聞き取りにくさを音声波形の観察で評価する簡易的な方法を用い、アクリル板に簡単な工事を施すことで聞き取りを改善できるか検討した。

アクリル板が音声伝達を妨げる原因として次の事が考

えられる。(1)空気とアクリル樹脂では音響インピーダンスが大きく異なるため、音エネルギーの多くはアクリル板表面で反射されてしまう。具体的に反射率を推定すると空気の音響インピーダンスはおよそ  $4 \times 10^2 \text{ kg/m}^2 \cdot \text{s}$  であり、アクリル樹脂の音響インピーダンスが  $3.6 \times 10^6 \text{ kg/m}^2 \cdot \text{s}$  という報告<sup>2)</sup>を基にすれば反射率99.98%となる。つまり、話者から直進的に発せられた音声のほとんどはアクリル板で反射されてしまい聴取者にはほぼ直接には届かない。(2)そのため聴取者に届く音声のほとんどはアクリル板の縁を回折して届くものと机の天板や部屋の壁や天井などで反響したものとなり、それらが聴取者の耳元で干渉して合成された音となる。(3)音の回折の点から見るとヒトの可聴域20 Hzから20 kHzの音は音速340 m/sとすれば波長がおおよそ17 mから17 mmとなる。特に音声の聞き分けに重要とされる300 Hzから4 kHzの周波数帯域<sup>3)</sup>に限っても1.1 mから85 mmとなり、低音は回折により到達しやすいが高音になるとアクリル板の陰になる場所には到達しにくいことになる。加えてアクリル板や室内での反射により様々な経路長を取った音が合成されるため位相情報が乱されることになる。これらにより音声の聞き取りに重要とされる時間微細構造<sup>4,7)</sup>の乱れ、つまり本来必要な波形特徴量の減少が予想される。

これらの予想を特別な機材を使わずに実験的に確かめるとアクリル板の飛沫防止効果を保ったまま音声伝

\*1 工学部医用生体工学科学部生

\*2 工学部医工学科教授

達を改善することを目指して本研究では次の実験を行った。通常の実験室内の実験机の上に設置したアクリル板の有無によって、到達する音声波形の時間微細構造がどの程度変化するかをスピーカーで再生した音声をマイクで記録することにより観察した。次にアクリル板のほぼ中央部に視認性を妨げない程度の開口部(窓)を設け、そこへ空気の流れを遮断しかつ音の透過を妨げにくい薄いフィルムを貼ることで音声の時間微細構造がどの程度保たれるのか検討した。その結果、アクリル板設置により音声波形の高周波成分の振幅と山谷数の減少が見られ、それがアクリル板に開口部を設けることで一定程度減少が抑えられること、そしてその開口部を厚さ 0.011 mm のポリ塩化ビニリデンフィルムで密閉しても音声波形への影響は少ないことが確認された。この方法を用いればコストをほとんどかけずに音声伝達に必要な時間微細構造の状態を簡易評価できる。それに加え、簡単で安価な加工法によりアクリル板のパーティションとしての利点を保ちつつ音声伝達も改善できることが示された。

## 2. 実験方法

通常の実験室内の実験机上に幅 800 mm 高さ 600 mm 厚さ 5 mm のアクリル板 (Aurako 製) を設置した。アクリルパーティションを介した人の会話を想定し、アクリル板を挟むようにして各々 15 cm 離れた位置に USB スピーカー (サンワサプライ製, MM-SPU9BK) と全指向性バックエレクトレットコンデンサー型モノラルマイクロフォン (オーディオテクニカ製, AT9931PC) を設置した。人の口と耳の位置に近くなるようにスピーカーの中心高さは机上 27 cm, マイクの中心高さは 34 cm とした。アクリル板に窓を設定する場合は、会話者同士の視認性を損なわないことを考慮して眼の高さより下になるようにした。具体的にはアクリル板の左右の中心かつ下端から 22.5 cm を中心とした幅 10 cm 高さ 5 cm の窓とした (Fig. 1)。窓に貼るフィルムとしてはポリ塩化ビニリデン (PVDC: polyvinylidene chloride) フィルム (サランラップ®, 旭化成製, 厚さ 0.011 mm) を使用した。なお、使用した PVDC フィルムは水や多くのガスを通さず自己接着性 (密着性) を持つ<sup>8-10)</sup>。アクリル板にも接着剤などを使うことなく密着させることができた。

音声は PC (Dell 社 Dimension E520) に保存した NHK ラジオニュースの音声ファイル (サンプリング周波数 48 kHz 約 2.5 分間の MP3 ファイル) を再生することで行った。音の再生と録音は波形編集ソフト (Audacity® Ver. 2.0.5) を用い、一連の実験中にスピーカーの音量とマイクのゲインは一定とした。実験室の空調は作動しているが、周囲で会話はされていない条件で測定を行った。マイクで録音された音データは波形編集ソフト (Audacity® Ver. 3.2.5) を用いて波形表示し、そのうちの典型的であった部分を Sample Data Export 機能により数値データとして書き出し、それを表計算ソフト (Microsoft 社, Excel)

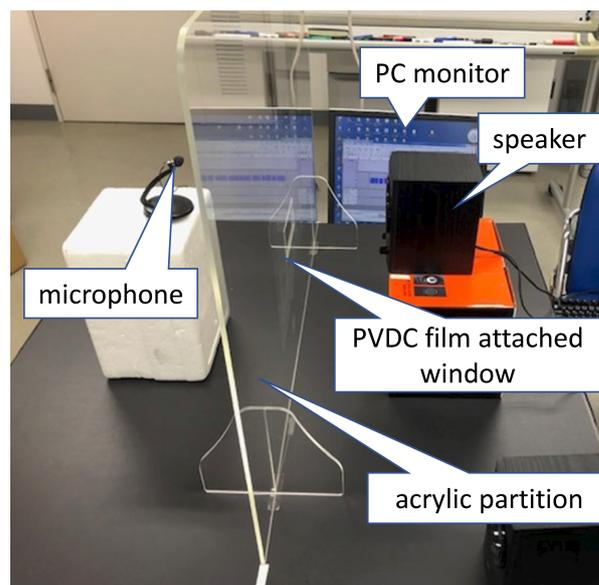


Fig. 1 Photograph of experimental set up.

によりグラフ表示した。なお、録音した音データは一切の加工をせずに波形観察に使用した。

なお、著者らはこの研究に関連したいかなる利益相反も無い。

## 3. 結果と考察

### 3.1 アクリル板設置による音声波形の変化

アクリル板無しでの録音を対照としてアクリル板を設置した場合の録音を聴き比べると音声全体がこもって (くぐもって) ぼやけた感じになり、いくつかの単語が聞き取りにくくなったという印象を著者らは受けた。ただし、完全に聞き取れない単語があるということでは無かった。特に母音については聞き取りが難しいとは感じられなかった。一方いくつかの子音、具体的には、無声破裂音 (/t/, /k/, /p/ など) と有声破裂音 (/d/, /g/, /b/ など), 無声破擦音 (/chi/, /tsu/ など) で聞き取りが難しくなる傾向を感じた。無声破裂音を子音とする /ta/ と有声破裂音を子音とする /ba/, そして無声破擦音である /chi/ を含む文である「～の立場を～」の部分の音声波形を示したのが Fig. 2 である。Fig. 2A にはこの部分全体の波形, Fig. 2B には注目した子音部分の拡大波形を示している。アクリル板無しの状態と有りの状態を比較するとアクリル板によって波形の高周波成分が減弱していることが波形全体から読み取れる。また、子音部分に注目すると /t/ の特徴である図中 0.15 s 付近の高周波の山谷がアクリル板により振幅が小さくなり数も減ることが見られた。より具体的には、0.14 s から 0.16 s の間に見られる先行する /no/ の母音の反響と考えられる約 5 周期の基本振動に、複合する形で /t/ の成分である高周波の山谷が主なものだけでも 30 以上、しかもそこにはより高周波の成分も伴って見られたのに対し、アクリル板によって /t/ 成分の振幅は大きく減り 20 程度のなめらかな山谷になっていた。/chi/

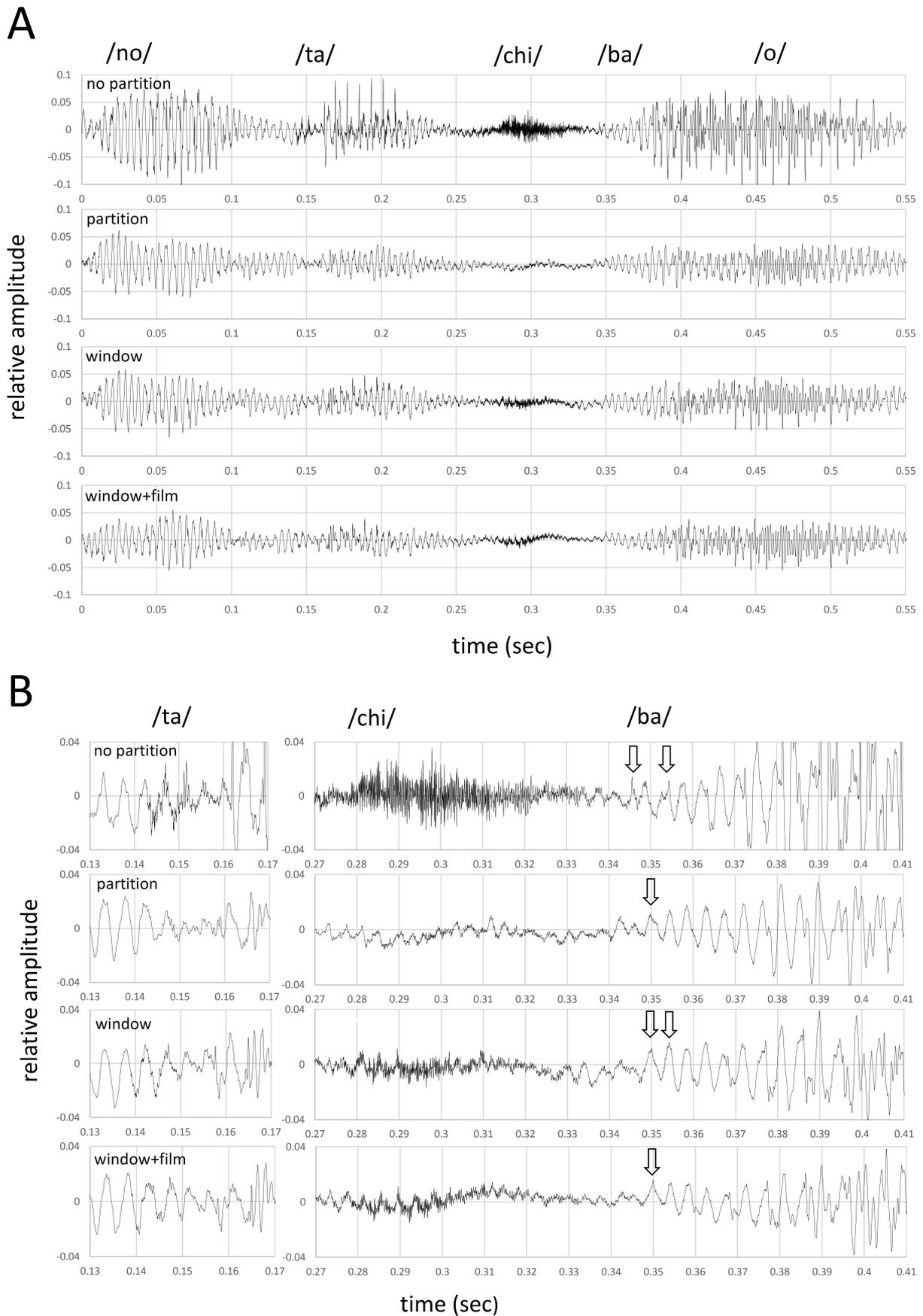


Fig. 2 Waveforms of “no-ta-chi-ba-o”. Comparison of control (no partition), partition, partition with window, and partition with PVDC film attached window. Whole waveforms (A), and enlarged regions of consonants (B).

については高周波の摩擦音に相当する波形がアクリル板によってほぼ見えなくなり、先行する/ta/の母音部分との境界をはっきりと特定するのが難しくなった。また、/ba/については立ち上がりの破裂音部分 (Fig. 2B 図中の矢印) がアクリル板無しでははっきり見られるものの、アクリル板有りでははっきりしなくなっていた。なお、これらそれぞれの音の部分だけを取り出して再生してみたところ、アクリル板有りでは/ta/が/pa/に聞こえたり、/chi/の認識が難しかったり、/ba/が/wa/のようにも聞こえたというのが著者らの印象だった。ただし、一連で再生すると「～の立場を～」と聞き取ることは可能だった。これは先行する音によって聴覚の順応状態が変わるために音の周波数帯域が変化することに聴覚が敏感であることと、これまでの言語学習によって不明瞭な部分を文脈で補って認識する<sup>3)</sup>ためと推測される。

母音の波形について見るとこちらでも高周波成分の減弱が見られた。しかし、母音それぞれの特徴を大きく決める第1フォルマントと第2フォルマントの成分は日本語では3.5 kHzまでに収まっていると報告されており<sup>11)</sup>、母音の識別に重要なこれらの波形成分はほぼ保たれていることがアクリル板有無での波形の直接観察比較とスペクトル解析の面(data not shown)から確認された。また、母音は継続時間が長い聴覚の神経処理としては一連

の繰返し周期内の時間微細構造を集積して認識につなげることになり、経過時間が短い子音成分より認識が容易であると考えられる。また、再生に使用した元の音声ファイルの波形と比較したところ実験環境下での反響によると思われる母音成分の継続時間の延長が見られ、これも母音の認識に有利にはたらくことが予想される。一方、これによって無声子音の前後に入っている無音部分に母音の反響が重なってしまうことは、子音の認識を低下させる一因になっていることが考えられる。

### 3.2 音声波形に対する窓とフィルムの効果

アクリル板に窓を設け音声を記録した場合の波形データとその窓にPVDCフィルムを貼って記録した時の波形を Fig. 2A,B の下2段に示した。これらに示されるように、アクリル板で少なくなった先行する母音の基本振動に重複した/t/の高周波成分の山谷の数と振幅が、窓によって部分的に回復されており、PVDCフィルム貼付でもその傾向は保たれていた。/chi/では高周波成分の振幅がアクリル板で1/6程度になっていたものが、窓によって1/3程度にまで回復し、PVDCフィルム貼付によっても1/4程度の振幅が見られた。これによって破擦音としての開始時と摩擦音が認識しやすくなると考えられる。

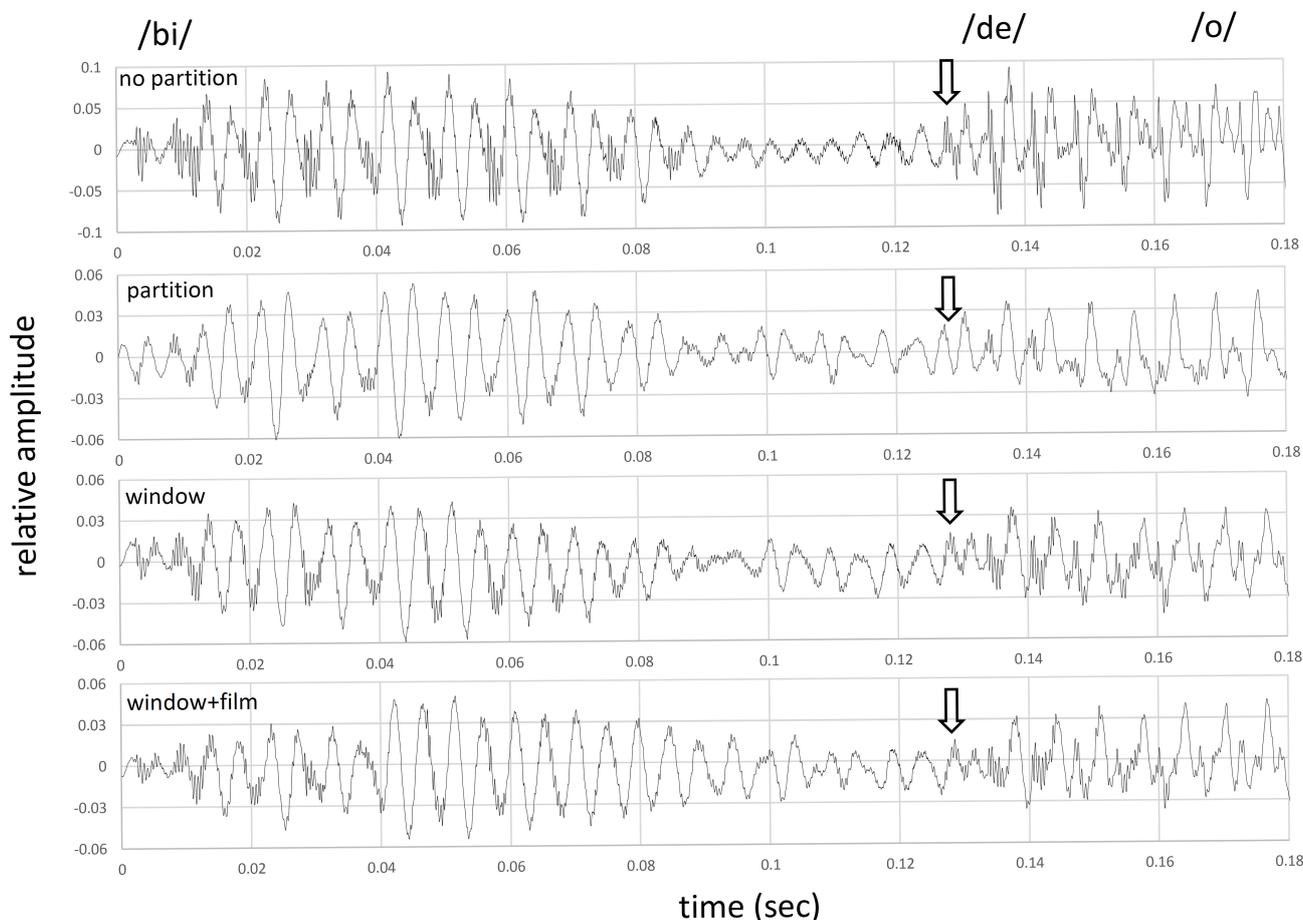


Fig. 3 Waveforms of “bi-de-o”. Comparison of control (no partition), partition, partition with window, and partition with PVDC film attached window.

/ba/については、/chi/の高周波成分が収まってからの/ba/への移り変わりがアクリル板ではっきりしなくなっていたものが、窓によってははっきりするようになり、/ba/の立ち上がり部分 (Fig. 2B 図中の矢印) の認識がしやすくなると考えられた。この傾向は PVDC フィルム貼付でも保たれていた。

一方、母音について見ると窓によって高周波成分が回復する傾向は見られるものの、全体的な振幅はアクリル板無し程度までは回復せず、窓無しのアクリル板有りとほぼ同じレベルであった。このことは音量的にはアクリル板設置による減少を今回の窓によっては回復させられていないことを示している。

Fig. 3 は「ビデオ」という音声の波形を示したものである。アクリル板により有声破裂子音である/b/の立ち上がり部分の高周波成分の振幅が减弱されている。また/de/の立ち上がり開始部 (Fig. 3 図中の矢印) が不明確になっている。このためか、アクリル板有りの音声を聴くと「ビオ」または「ビヨ」のようにも聞こえるという印象を著者らは持った。これは、アクリル板の設置により/d/の開始部分が不鮮明になることと母音/e/から/o/に移移する部分も変化が不鮮明なために起きている可能性が

考えられる。日本語の母音の第1フォルマントと第2フォルマントは、/e/では概ね 450 Hz と 2000 Hz, /o/では概ね 430 Hz と 650 Hz とされており<sup>12)</sup>, /e/の第2フォルマントの認識がアクリル板によって難しくなったために第1フォルマントの周波数が比較的近い/e/と/o/が区別しにくくなった可能性が考えられる。なお他の母音では、/a/が概ね 790 Hz と 1210 Hz, /i/が概ね 280 Hz と 2210 Hz, /u/が概ね 310 Hz と 1100 Hz とされており<sup>12)</sup>, これらは/e/と/o/とは第1フォルマントが大きく異なることからこの推測は妥当であると考えられる。アクリル板に窓を設け PVDC フィルム貼付したところ、/b/の立ち上りの高周波波形がはっきりすること、/de/の開始部 (Fig. 3 図中の矢印) がはっきりすること、/e/と/o/の波形も高周波成分が多かった状態から少ない状態へと波形が明確に変化するのが見られた。これらが「ビデオ」の聞き取りを明瞭にすることに繋がっているものと思われる。

Fig. 4 は/ka/の波形全体を左の列に、/k/部分の拡大を右の列に示したものである。なお、縦軸も右の列で拡大されている。この図を見ると無声破裂子音である/k/の特徴的な振動は、およそ 10 ms あたり 16 個の山谷つまり約 1.6 kHz であると推定される。この振動が母音/a/に先行し

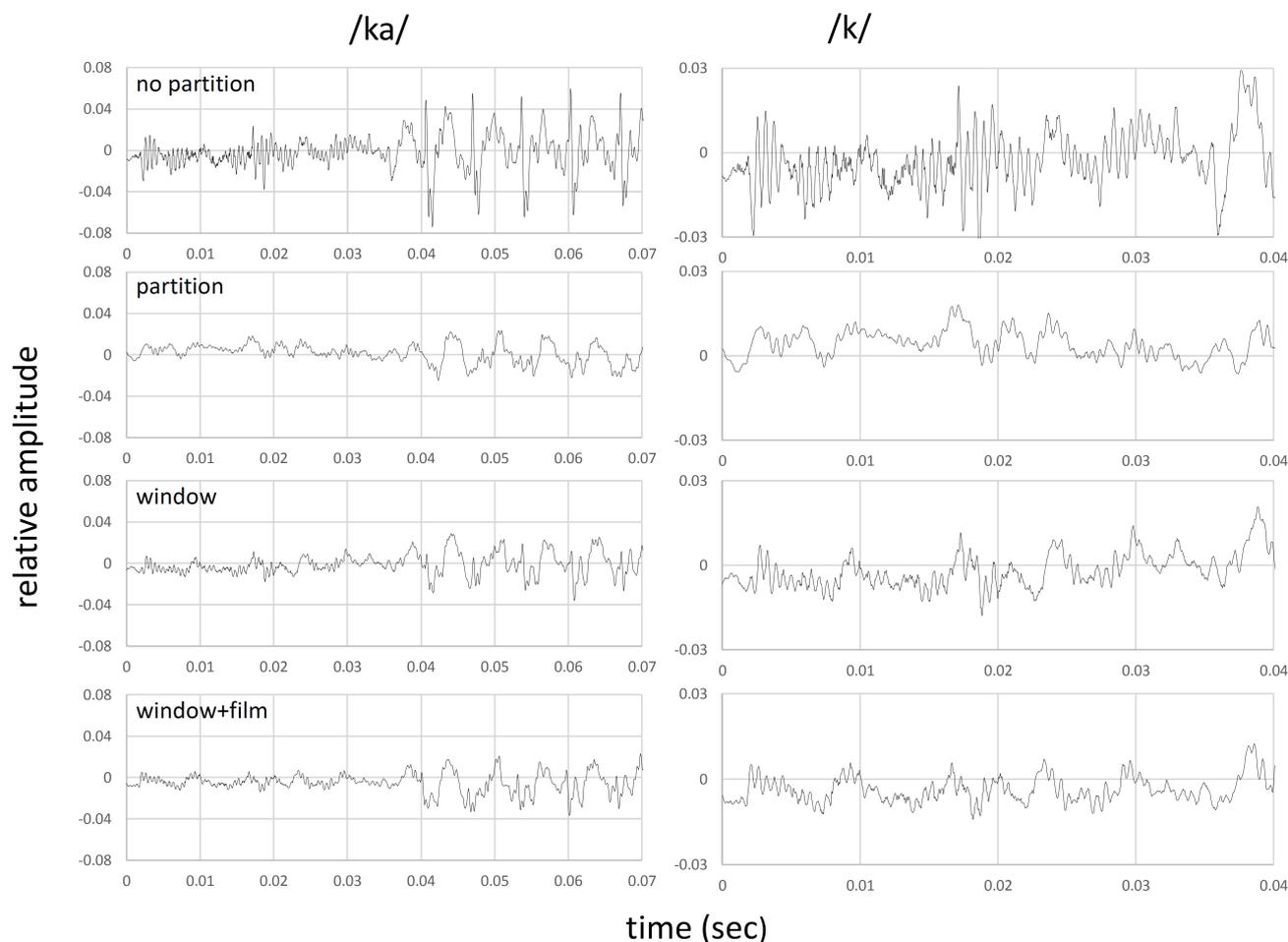


Fig. 4 Waveforms of “ka”. Comparison of control (no partition), partition, partition with window, and partition with PVDC film attached window. Right column indicates enlarged region of consonant /k/.

て入ることが/kɑ/の認識に必要であると思われる。この子音部がアクリル板によって振幅が大きく抑制されており、山谷の数がはっきりしにくい時間領域が増えている。これらが/k/認識が難しくなる要因になると考えられる。一方、窓に PVDC フィルム貼付したものでは/k/部分の振幅に部分的な回復が見られており、認識しやすくなっていると考えられる。なお、実際に著者らが音声を聴いた印象では、/k/の前がほぼ無音の状態であればアクリル板を介した音声でも/k/を認識するのは容易であるが、/k/に先行して母音がありその反響が長引くと/k/の認識が難しくなるというものだった。例えば、単語中の/kɑ/の部分だけを聴くと/dka/のよう感じられたり、/ku/が/kbu/または/fu/のように感じられた例や、/ko/が/ho/や/po/に感じられた例もあった。なお波形データは示さないが、ここまで示した以外の子音でアクリル板により認識が不明瞭になると著者らに感じられた例としては、/gi/が/bi/や/ni/または/i/と区別するのが難しくなった例、/ze/が/de/と区別しにくくなった例、/hi/が/pi/と区別しにくくなった例、/tsu/が/zu/と区別しにくくなった例などが感じられた。これらの子音の認識の明瞭化に対しても、窓に PVDC フィルム貼付した場合に一定の改善効果があると感じられた。

#### 4. おわりに

今回の測定方法は厳密な音響学的な計測ではない。しかし、単に一連の音声データを FFT により全体的なパワースペクトルを見ただけでは音声の認識に関わる要素である波形 1 周期内の山谷の構造、その経時変化、そして特定の波形の継続時間といった分析は十分にはできない。今回目的としたのは一般的に入手可能な機材で再現実験が可能であり、音波形の直接観察といった簡易的な方法であっても音声伝達における聞き分けに重要となる子音成分の修飾や継続時間の変化といった時間微細構造に関わる要素を定性的に評価できることである。そして、その目的は達成されたと考えている。我々の研究室では既にクラリネットの発音や音色の知覚に関わる波形要素の見積もり<sup>13)</sup>や、アルトリコーダーで吹鳴された重音の知覚に関わる波形要素の提案<sup>14)</sup>において、今回同様の波形観察方法の有用性を示している。従って、本研究で行った測定・観察方法は、今後それぞれ具体的に異なるアクリルパーティションの設置環境においても、音声伝達の良否を定性的に評価するための実用的な方法として提案できたと考えている。これに加えて、本研究では次の事も示すことができた。アクリル板は切削といった加工性も高く窓を設けることは難しくない。PVDC フィルムは薄くても強度があり破れにくく音声の透過性が良い一方、飛沫の透過は防ぐ。PVDC フィルムはアクリル板にも密着し、平面的に貼れば窓部分を介しての視認性も妨げにくい。しかも安価で入手しやすく、必要に応じて簡単に張り替えられるので衛生的に使用できる。以上により、

アクリルパーティションに窓を設け PVDC フィルムを貼るという方法は、聞き取りを簡易的に改善する方法として有効かつ実用的であると言えるであろう。

実際の会話の状況では様々な環境雑音や、他者の会話が聞こえていることも多いであろうから、今回の実験条件より更に聞き分けが難しい場面が多くなると考えられる。アクリル板を使用していると聴取者の後ろから届いてくる音がアクリル板の反射により大きく聞こえることも体験される。そういった状況下では人は両耳機能を用いて妨害音の中から聞きたい音声を音源定位によって聴取する事が可能なことが知られている<sup>3)</sup>。しかし、アクリル板によって正面からの音が遮られてしまうと、視覚情報から判断される正面である話者の位置と音源定位能力による音声源の位置が一致なくなってしまう。このような場合でも、PVDC フィルム貼付の窓があれば、そこを通して直接音が聴取者に届くことになり、より話者の音声を選択的に認識することに有利になるものと予想される。そのため今後は、環境雑音や他者の会話が聞こえているような状況下で、被検者による実地試験を行い、アクリル板に PVDC フィルム貼付の窓を設けることの有効性を様々な実環境下で具体的に評価する必要がある。なおそれと並行して、より聞き取りに優れたアクリルパーティションとして実用化するためには、聞き分けに重要な子音成分を効果的に透過させられる窓の大きさや形状そして位置について検討することも必要であろう。高齢者では老化により高周波域の聴力が低下し、同時に音声分析する能力も低下することが知られている<sup>3)</sup>。現在多く使用されているアクリルパーティションでは音声コミュニケーションに難しさを感じている人々がいることを忘れずに改良を進めていく必要があるであろう。

#### 参考文献

- 1) 青木佐奈枝, 小野聡士, 福井晴那, 川嶋真紀子: COVID-19 感染防止対策が心理支援にもたらした問題とその対応, 心理学研究, Vol. 93, No.4, pp.359-365 (2022).
- 2) 久本誠一, 樋口雅俊, 田中均, 藤井尚子, 合田真理子, 上津遊恭平: 水槽を併用した超音波エコー法による体肢断面観察法の検討, パイオメカニズム学会誌, Vol. 34, No.1, pp.68-72 (2010).
- 3) 廣谷定男(編著): 聞くと話すの脳科学 (日本音響学会編, 音響サイエンスシリーズ 17), コロナ社 (2017).
- 4) M.S. Lewicki: Efficient coding of natural sounds, Nature Neurosci., Vol.5, No.4, pp.356-363 (2002).
- 5) E.C. Smith and M.S. Lewicki: Efficient auditory coding, Nature, Vol.439, pp.978-982 (2006)
- 6) B. C. J. Moore: The role of temporal fine structure processing in pitch perception, masking, and speech perception of normal-hearing and hearing-impaired people, J. Assoc. Res. Otolaryngol., Vol.9, pp.399-406

- (2008).
- 7) J.N. Oppenheim and M.O. Magnasco: Human time-frequency acuity beats the fourier uncertainty principle, *Physical Review Letters*, Vol.110, pp. 044301(1-5) (2013).
  - 8) 板谷実: プラスチックフィルムによる食品の包装, *調理科学*, Vol.3, No.2, pp.73-78 (1970).
  - 9) 市川和昭: 食品包装の研究 (第1報) プラスチック包装材料についての一考察, *名古屋文理短期大学紀要*, Vol.18, pp.101-107 (1993).
  - 10) 中尾卓: サランラップ®の製膜とフィルム特性, *化学と教育*, Vol.58, No.12, pp.558-561 (2010).
  - 11) T. Hirahara and R. Akahane-Yamada: Acoustic characteristics of Japanese vowels, *Proc. ICA2004*, pp.3287-3290 (2004).
  - 12) 加藤重弘, 安藤智子: 基礎から学ぶ音声学講義, 研究社, pp.223 (2016).
  - 13) 堀越哲郎, 藤原正明: クラリネットマウスピースのフェイシング部の微小形状変化が吹鳴音波形に与える影響, *東海大学紀要工学部*, Vol.62, No.1, pp.35-42 (2022).
  - 14) 堀越哲郎, 藤原正明: アルトリコーダーで得られた重音波形の時間構造に見られた特徴, *東海大学紀要工学部*, Vol.59, No.2, pp.51-58 (2019).