

口蓋裂患者に対するテレプラクティスの音声評価

田尻, 姿穂

<https://hdl.handle.net/2324/7363717>

出版情報 : Kyushu University, 2024, 博士 (歯学), 課程博士
バージョン :
権利関係 :



口蓋裂患者に対するテレプラクティスの音声評価

田尻 姿穂

九州大学大学院歯学府
口腔顎顔面病態学講座 口腔顎顔面外科学分野

指導： 森山 雅文 教授

九州大学大学院歯学府
口腔顎顔面病態学講座 口腔顎顔面外科学分野

本研究の内容の一部は下記の学術雑誌に掲載されたものである。

Acoustic evaluation of voice signal distortion by videoconferencing platforms and devices used in telepractice for cleft palate

Shiho Tajiri, Shunsuke Hidaka, Shuhei Takehisa, Sachiyo Hasegawa, Yukiko Ohyama, Tomohiro Yamada

Congenital Anomalies. 2024; 64(6): 242-253

略語一覧

VCP: Videoconferencing platform (テレビ会議システム)

CC: Cross-correlation (相関係数)

STFT: Short-time Fourier transform (短時間フーリエ変換)

DFT: Discrete Fourier transform (離散フーリエ変換)

DTW: Dynamic time warping (動的時間短縮法)

ED: Euclidean distance (ユークリッド距離)

LSD: Log spectral distance (対数スペクトル歪み)

PA: Palatalized articulation (口蓋化構音)

LA: Lateral articulation (側音化構音)

GS: Glottal stop (声門破裂音)

VOT: Voice onset time (有声開始時間)

目次	3
要旨	4
緒言	5
研究 1. テレプラクティスにおける音声信号歪みの音響学的評価	
目的、材料と方法	7
結果	15
小括	21
研究 2. テレプラクティスにおける言語聴覚士の構音評価に与える影響	
目的、材料と方法	23
結果	24
小括	29
考察	31
結語	34
謝辞	35
参考文献	36

要旨

口蓋裂患者では約 50%に構音障害が認められ、代償性構音は言語聴覚士による言語療法が行われる。言語療法は言語聴覚士による聴覚知覚的識別に基づいており、誤った調音位置や調音法を修正し、適切な音声機能確立することを目的とする。近年、COVID-19 の流行により遠隔医療の需要が高まっており、口蓋裂患者に対するテレプラクティスに関しても、その有用性や有効性が多数報告された。しかし、これまでの報告は、短期・中期的な治療成績、臨床家や患者の満足度調査に基づくものが多く、テレプラクティスの音声評価に関する報告はほとんどない。VCP を介した音声は、音声信号歪みが生じ、患者の認識や臨床医の評価に影響する可能性がある。

そこで本研究では、4 つの VCP (Zoom®, Cisco Webex®, Skype®, Google Meet®) と 5 つのデバイス (MacBook®, Windows®, iPad®, Android®, iPhone®) の組み合わせにおける音声信号歪みを客観指標である LSD を用いて音響学的に定量評価した。さらに口蓋裂言語のうち頻度の高い PA、LA について、VCP を介した音声の音声信号歪みと言語聴覚士による構音判定との関係を検証した。これらの検証により、VCP とデバイス間での音声信号歪みの程度、音声信号歪みの小さい音素、VCP の構音判定に及ぼす影響を明らかにすることとした。

LSD は小さいほど音声信号歪みが小さいことを示すが、正常音声サンプル ATR503 文を用いた検証では、VCP 間の平均 LSD 値は、Zoom® が最も大きく、Google Meet® が最も小さかった。デバイス間の平均 LSD 値は、iPhone® が最も大きく、MacBook® が最も小さかった。子音は無声、歯茎硬口蓋音、摩擦音/破擦音で有意に平均 LSD 値が大きく、有声、両唇音/唇軟口蓋音、鼻音/破裂音で有意に平均 LSD 値が小さかった。口蓋裂言語である PA、LA を用いた検証では、オリジナル音声と VCP を介した音声の言語聴覚士による構音判定は、全ての VCP においてかなりの一致を示した。Zoom® と Cisco Webex® が、Skype® と Google Meet® と比較し、有意に平均 LSD 値が大きかったが、平均 LSD 値 7.15~10.30 dB の範囲では、全ての VCP 間で言語聴覚士による構音判定の一致率に有意差はなかった。

以上の結果より、テレプラクティスにおいて VCP やデバイスの種類によって歪みの程度に差があり、適切な VCP やデバイスの選択が必要である可能性が示唆された。PA、LA における検証では、VCP による音声信号歪みは生じるものの、言語聴覚士による聴覚知覚的音声評価の精度には大きく影響しないことが明らかになり、テレプラクティスの有用性が示唆された。

緒言

口蓋裂は最も一般的な先天異常の1つである。国際的な有病率は10,000人あたり7.94人で、特に日本では10,000人あたり19.05人と高い[1]。口蓋裂患者は鼻咽腔閉鎖機能の獲得を目的に口蓋形成術が行われ[2]、早期の言語介入が必要である。早期言語介入の最も重要な目的は、子音の数を増やすことによって言語発達を助けることである[3]。口蓋形成術後、口蓋裂患者の約50%で構音障害が認められるといわれている[4]。鼻咽腔閉鎖機能不全や口腔鼻腔瘻孔によって生じる構音障害は外科的介入を必要とするのに対し、代償性構音は言語聴覚士による言語療法によって改善される。現在、口蓋裂患者に対する言語療法の評価と診断は、言語聴覚士による聴覚知覚的識別に基づいている[5]。言語聴覚士による評価では、音声エラーを鑑別し、介入に適した対象音を決定する。また、代償性構音のある口蓋裂患者の言語療法は、誤った調音位置や調音法を修正し、適切な音声機能を確立することを目的とする[6,7]。

近年、COVID-19の流行により、遠隔医療の需要が高まっている[8]。音声言語病理学分野では、アメリカ言語聴覚士団体であるAmerican Speech-Language Hearing Associationが、遠隔コミュニケーションやインターネット技術を用いたスクリーニング、評価、介入、コンサルトのための臨床家と患者の遠隔での言語病理学サービスとしてテレプラクティスを提唱している[9]。COVID-19は飛沫や物理的接触を介して感染する可能性があり[10,11]、COVID-19の流行下において、テレプラクティスは有用であると多数報告された[12,13]。また、遠隔医療は地方に住む患者家族の移動負担の軽減、国際的な医療教育、専門医の不足による地域格差の是正にも有用で、口蓋裂患者に対する言語療法に関してもその有用性や有効性が多数報告された[14-16]。しかし、これまでの口蓋裂患者へのテレプラクティスに関する報告は、短期・中期的な治療成績や臨床家や患者の満足度調査に基づくものが多く、音声評価に関する報告はほとんどない。

VCPを使用したテレプラクティスでは、送信デバイスのマイク、VCPを介した音声データのエンコード、ノイズ、受信デバイスのスピーカー、音声強調アルゴリズムやノイズキャンセリング、自動音量調整機能などの様々なVCPの機能[17-19]、インターネット帯域[20,21]などの要因によって音声信号歪みが生じる可能性がある。音声信号歪みは、患者の知覚や発話を正確に再現することや、臨床医の評価に影響を与える可能性がある。

そこで、研究1ではテレプラクティスによる音声信号歪みについて音響学的に検証し、研

究2では主観評価に及ぼす影響を検証することで、VCPを介した音声の音声信号歪みの程度、音声信号歪みを受けやすい音素、構音判定に及ぼす影響を検討した。

研究 1

1. 目的

VCP を使用したテレプラクティスでは、音声データのエンコード、VCP 搭載の機能、受信デバイスのマイクやスピーカーなどの要因で、音声信号歪みが生じる可能性がある。しかし、音響学的な観点から、テレプラクティスに用いる適切な VCP やデバイスは決まっておらず、これまで口蓋裂患者に対して行われたテレプラクティスに関する報告は、報告者によって使用している VCP やデバイスが異なる。また、口蓋裂患者の言語訓練では、音声エラーを鑑別し、介入に適した対象音を決定することが重要であるが、テレプラクティスでは音声信号歪の受けやすい音素は明らかになっていない。そのため、音声信号歪みは言語聴覚士の評価や治療効果に影響する可能性がある。そこで本研究の目的は、さまざまな VCP とデバイスの組み合わせを用いて音声信号歪みを比較し、VCP とデバイス間での音声信号歪みの程度の差および音声信号歪みの小さい音素を音響学的に定量評価することである。

2. 材料および方法

2-1. 音声データ

音声データベースである ATR503 を用いた。ATR503 は音声学的にバランスのとれたテキストコーパスであり、日本語に含まれる可能性のある全ての 402 の 2 連鎖音素と、前後の影響を受けやすい 223 の 3 連鎖音素、合計 625 がバランスよく含まれている文章である [22,23]。ATR503 は 10 セット (A~J) で構成されているが、今回の検証で用いたセット A は、現在および過去に構音や聴覚障害がない 1 人の日本人男性によって収録されたものである。今回の検証では口蓋裂言語による音声信号歪みの影響を排除するために、正常音声サンプルを用いた。また、日本語の全子音のフォルマント周波数に男女間で有意差はないとされているため [24]、男性の単一音声を選択した。音声データは防音室で無指向性マイクロホンアンプを用いて録音され、サンプリング周波数は 48 kHz、ビット深度は 16 bit であった。

2-2. VCP とデバイス

4 つの VCP—Zoom® (ver. 5.10.4), Cisco Webex® (ver. 42.5.0/42.5.1), Skype® (ver. 8.83), and Google Meet® (ver. 89.0.0) を用いた。本邦における「オンライン診療の適切な実施に関する指針 (厚生労働省)」では、セキュリティ対策として、信頼できる機関が発行したサ

ーバー証明書による通信の暗号化 (Transport Layer Security 1.2 以上) を実施することが明記されており [25]、これらの VCP はこのセキュリティーレベルに適合している。VCP は PC 用の Google Meet®を除き、ダウンロードしたアプリケーションを使用した。PC 用の Google Meet®はアプリケーションがなかったため、Google Chrome 上で実行した。各 VCP は、音量とノイズの自動調整が可能である。全てのデバイスが同じ条件で使用されるように、デフォルト設定、つまり音量自動調整機能、ノイズキャンセリング機能を用いた。Google Meet®無料版に関してはこれらの設定がなかったため選択できなかった。

5 つのデバイス—ラップトップ 2 種類; Apple MacBook Pro 13-inch M1 2020® (macOS Monterey 12.4), Microsoft Surface Go® (Windows 10), タブレット 1 種類; Apple iPad Air 第 4 世代® (iOS 15.5), スマートフォン 2 種類; Samsung Galaxy A21® (Android 12.0), Apple iPhone 8® (iOS 15.5)を用いた。これらは、本邦の OS の普及率とデバイスの保有率 (スマートフォン 86.8%、ラップトップ 70.1%、タブレット 38.7%) をもとに選定した [26]。

2-3. 機器のセットアップと音声信号

送信デバイスは前述のデバイスのいずれかを用い、受信用デバイスは MacBook Pro® (macOS Monterey 12.4) 1 種類のみを用いた。受信用デバイスは Python® (<https://www.python.org/>) を用い再生と録音に利用し、バーチャルオーディオインターフェイスを介することで伝送遅延をなくした。実験中は、オーディオ・インターフェース、ラウンドスピーカー、マイクアンプのゲインは固定した。Wi-Fi 接続によるパケットロスの影響を避けるため [27]、送受信デバイスは有線接続し、デバイスの変更ごとに Speedtest® (<https://www.speedtest.net/ja>) を用いてインターネット速度を測定した。音の空気伝達路に関わる機器は無響室に設置した。無響室内には騒音はなく、温度は 20.3°C~22.6°C、湿度は 27~28%に維持した。インターネット速度は実験全体を通じてダウンロード速度 150 Mbps、アップロード速度 100 Mbps 以上であった。不必要なネットワーク負荷を軽減するため、送信用デバイスのカメラとラウンドスピーカーはオフにした。

機器のセットアップを図 1 に示す。初めに、受信用デバイスからオーディオインターフェイスを介して無響室内のラウンドスピーカーに音声データを送信した。次に、ラウンドスピーカーを中心とした半径 0.5 m の同心円上に、マイクロホンをラウンドスピーカーと同じ高さに設置した。送信用デバイスがラップトップの場合は、机上に開いた状態で設置し、タブレットやスマートフォンの場合は、それらをタブレットスタンドに設置し、デバイスに

搭載されているマイクが左側になるように設置した。マイクアンプとオーディオインターフェースを介してマイクロホンに入力された音声は受信用デバイスで録音され、送信用デバイスに入力された音声は VCP を介して受信用デバイスであるラップトップに転送された。

2-4. 録音と保存

音声データの信号パワーを正規化し、最初にホワイトノイズをスピーカーから流した。実験中はマイク近くの音圧を 68.2 dB SPL に調整した。マイクロホンから受信した音声データを「入力波形」、VCP から受信した音声データを「出力波形」と定義した。シグナルチェーンを図 1 に示す。50 の文章はそれぞれ 2 回連続で再生した。1 回目は自動音量調整による予期せぬ音量変化を防ぐため音量調整用として利用し、2 回目は検証に用いるための入出力波形の収録に用いた。受信用デバイスの音声と、受信用デバイスに表示される対数振幅スペクトログラムによって、聴覚視覚的に異常がないかを確認しながら録音を行った。入出力波形はサンプリングレート 48 kHz、ビット深度は 32 bit で保存した。

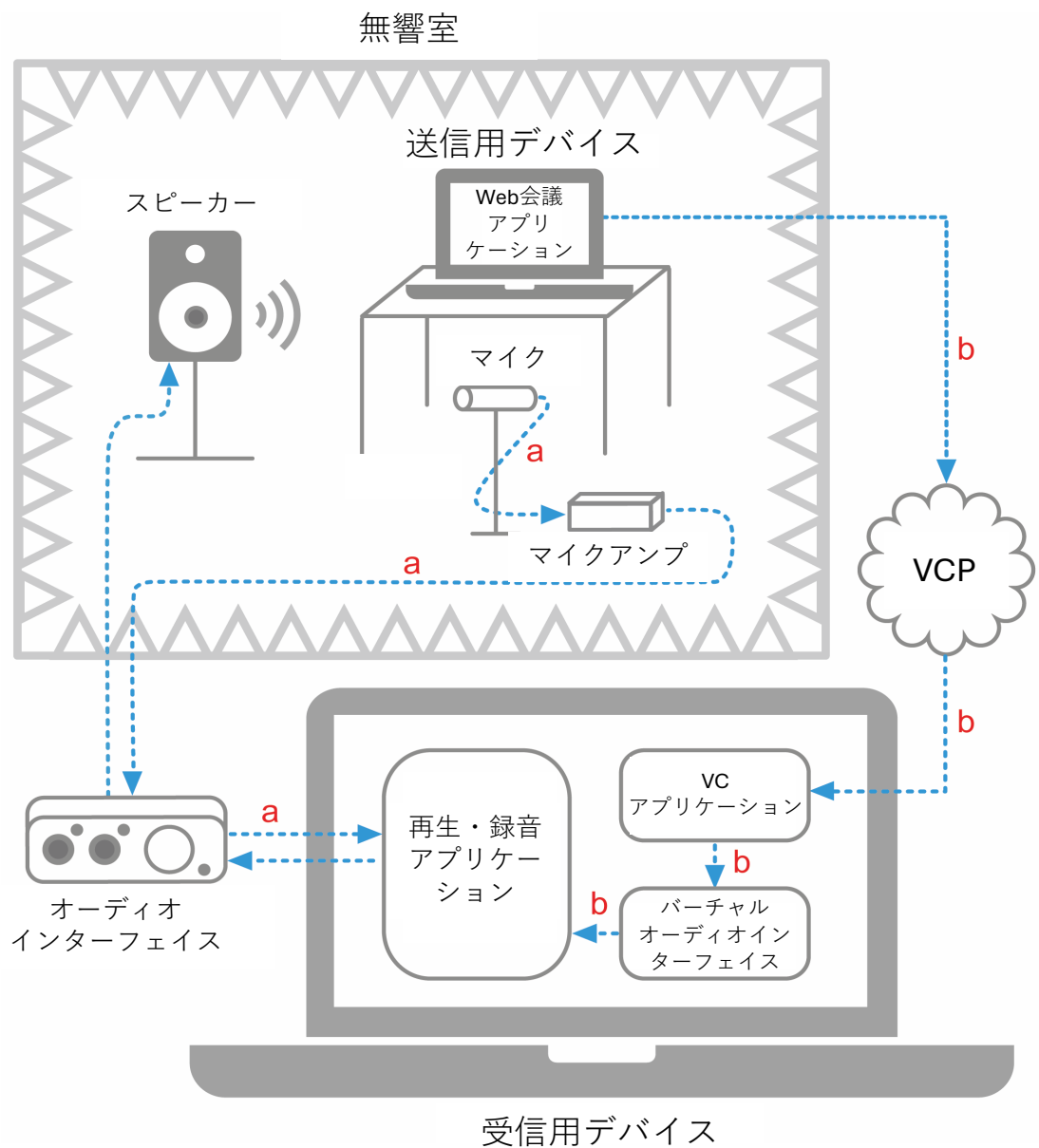


図 1 機器のセットアップとシグナルチェーン

a; マイクロホンから受信した音声データである「入力波形」の音声信号

b; VCP から受信した音声データである「出力波形」の音声信号

3. 音声データ分析

3-1. 声分析ソフトウェアを用いたアノテーションの付与

Praat® (<https://www.fon.hum.uva.nl/praat/>)を用いて、音声サンプルの文章を聴覚視覚的に音素に切り分け、IPA 国際記号をもとに音素を割り当てた [28]。例として「あらゆる」という音声サンプルを図 2A に示す。それぞれの音素を有声性 2 つ（無声と有声）、調音位置 7 つ（両唇音、歯茎音、歯茎硬口蓋音、硬口蓋音、軟口蓋音、唇軟口蓋音、声門音）、調音法 6 つ（破裂音、摩擦音、破擦音、はじき音、鼻音、接近音）に分類した。付与された音素の種類と数を表 1 に示す。

表 1 本研究に用いられた ATR503 文の IPA 国際音声記号の分類と数

	調音位置						
	両唇音	歯茎音	歯茎 口蓋音	口蓋音	軟口蓋 音	唇軟口 蓋音	声門音
調音法							
破裂音							
無声	p (115)	t (505)			k (700)		
有声	b (190)	d (280)			g (375)		
摩擦音							
無声	ɸ (80)	s (350)	ɕ (240)	ç (140)			h (125)
有声			z (80)				
破擦音							
無声		ts (130)	tɕ (145)				
有声		dz (130)	dʑ (140)				
はじき音							
有声		r (585)					
鼻音							
有声	m (385)	n (355)		ɲ (220)			
接近音							
有声				j (235)		w (185)	

3-2. 音声サンプルと入力波形の開始点の位置合わせ

音声サンプルに付与した各音素のアノテーションを入力波形に対応づけるため、音声サンプルと入力波形の時間差を最大相互関係 CC の時間によって決定し、入力波形の開始点を音声サンプルの開始点に一致させた。CC は以下のように定義した。(図 2B)

$$CC(k) = \sum_m S(m, n)I(m, n + k) \quad (1)$$

S は音声サンプル、 I は入力波形の振幅スペクトログラムを示す。 m は周波数、 n は時間のインデックス、 k はタイムラグを示す。振幅スペクトログラムは STFT により行うが、窓関数と離散フーリエ変換は 2400 サンプル(50 ms)、シフト長は 1 サンプル(1/48)とした。

3-3. 入出力波形の時間補正

入出力波形は VCP を介することで、遅延などの時間的ズレが生じている可能性がある。そこで、DTW を用いて入出力波形の時間補正を行った。DTW は音声認識における時間方向のアライメントに用いられる [29]。2 つの対数振幅スペクトル間の ED を DTW の距離関数として使用した。ED は以下のように定義した。(図 2C)

$$ED(i, j) = \sqrt{\sum_{m=1}^M (20 \log_{10} O(m, i) - 20 \log_{10} I(m, j))^2} \quad (2)$$

O は出力波形の振幅スペクトログラム、 M は直流成分からナイキスト周波数までの周波数ビンの数を示す。出力波形は VCP によって最低 8kHz に帯域制限されていたため、ED 計算のために入出力波形は 16 kHz にダウンサンプリングした。振幅スペクトログラムは STFT により行うが、窓関数と離散フーリエ変換は 800 サンプル(50 ms)、DFT ポイントは 4000 サンプル(250 ms)、シフト長は 200 サンプル(12.5 ms)とした。

3-4. LSD を用いた歪みの定量化

入出力波形の対数振幅スペクトログラム間の歪みを、LSD を用いてフレームごとに計算し、音声信号歪みを定量化した。LSD は 2 つの対数振幅スペクトル間の歪みの尺度である [30,31]。LSD の値が小さいほど歪みが小さいことを示す。LSD は以下の式を用いて定義した。

$$\text{LSD}(j) = \sqrt{\frac{1}{M - \tilde{M} + 1} \sum_{m=\tilde{M}}^M (20 \log_{10} \tilde{O}(m, j) - 20 \log_{10} I(m, j))^2} \quad (3)$$

$\tilde{O}$ は出力波形の時間補正後の振幅スペクトログラム、 $\tilde{M}$ は LSD 計算から 64 Hz 未満の周波数成分を除外するように設定した。入出力波形は 16 kHz にダウンサンプリングした。窓関数と離散フーリエ変換は 800 サンプル(50 ms)、DFT ポイントは 4000 サンプル(250 ms)、シフト長は 200 サンプル(12.5 ms)とした。

母音、子音、特殊モーラを含む全ての音素に対する音声信号歪みの程度を、LSD を用いて V C P、デバイス、それぞれの組み合わせごとに検証した。子音において、IPA 国際音声記号を用いてグループ間の歪みの程度を有声性、調音位置、調音法に関して比較した。

3-5. 統計分析

JMP(ver.16)を用いて行った。p< 0.05 を有意差とし、全ての音素を検証する際は一元化分散分析を行い、効果量 η^2 を求めた。その後、Tukey's 検定を用いて比較した。それぞれの音素に関しては、グループ間の比較を有声性は t 検定、調音位置と調音法は Tukey's 検定にて行い、Cohen's d を用いて効果量を計算した。

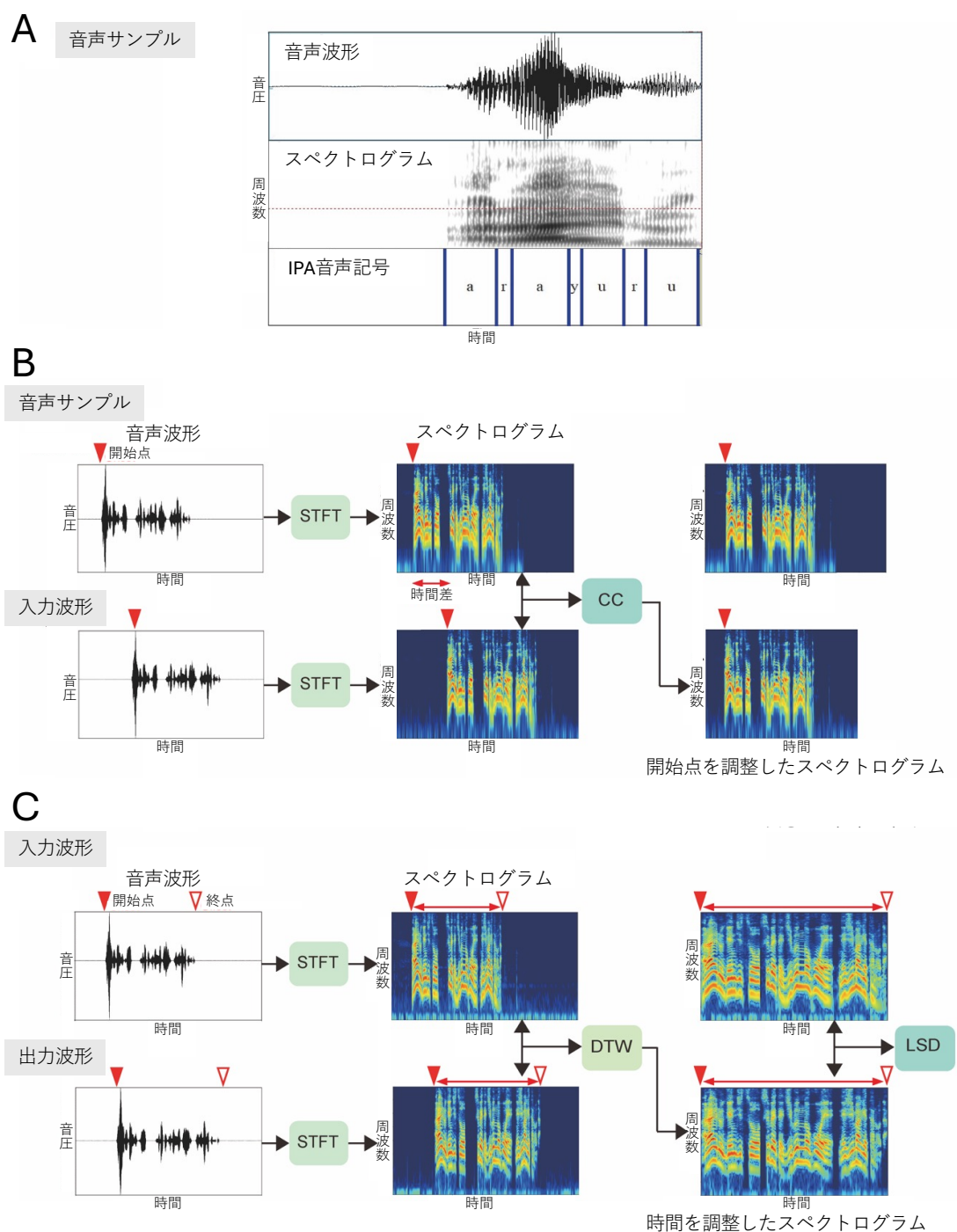


図 2 データ解析の概要

- (A) Praat®を用いて「あらゆる」という音声サンプルのアノテーションを付与した例。
- (B) 入力波形にアノテーションを付与するために、CC を用いて入力波形の開始点を音声サンプルの開始点にシフトさせた。
- (C) LSD を計算するために、DTW を用いて入力波形と出力波形の開始点から終点の位置合わせを行った。

4. 結果

4.1. 全ての音素における伝送条件の影響

一元化分散分析の結果、表 2 に示すように、VCP 間、デバイス間の両方で統計学的有意差を認めた。VCP は効果量中、デバイスは効果量小であった。VCP 間では、Zoom[®]が最も高い LSD 値を示し、続いて Cisco Webex[®]、Skype[®]、Google Meet[®]の順であった。(図 3A) Tukey's 検定の結果、全ての VCP 間で有意差を認めた。Zoom[®]と Google Meet[®]間では効果量大であった。デバイス間では、iPhone[®]が最も高い LSD 値を示し、続いて、Android[®]、iPad[®]、Windows[®]、MacBook[®]の順であった。Tukey's 検定の結果、全てのデバイス間で有意差を認めた。(図 3B) MacBook[®]と iPhone[®]間で効果量大であった。

表 2 VCP およびデバイス間の一元化分散分析とそれに続く Tukey's 検定の結果

	一元化分散分析	効果量	Tukey's 検定		効果量
	p 値	η^2	多重比較		Cohen's d
VCP	<0.001	0.093	Zoom	Cisco Webex	0.41
				Skype	0.72
				Google Meet	0.87
			Cisco	Skype	0.27
			Webex		
				Google Meet	0.38
			Skype	Google Meet	0.09
デバイス	<0.001	0.083	MacBook	Windows	0.49
				iPad	0.58
				Android	0.69
				iPhone	0.84
			Windows	iPad	0.14
				Android	0.24
				iPhone	0.43
			iPad	Android	0.10
				iPhone	0.29
			Android	iPhone	0.20

η^2 効果量: 0.01 =小, 0.09 =中, 0.25 =大

Cohen's d 効果量: 0.2 =小, 0.5 =中, 0.8 =大

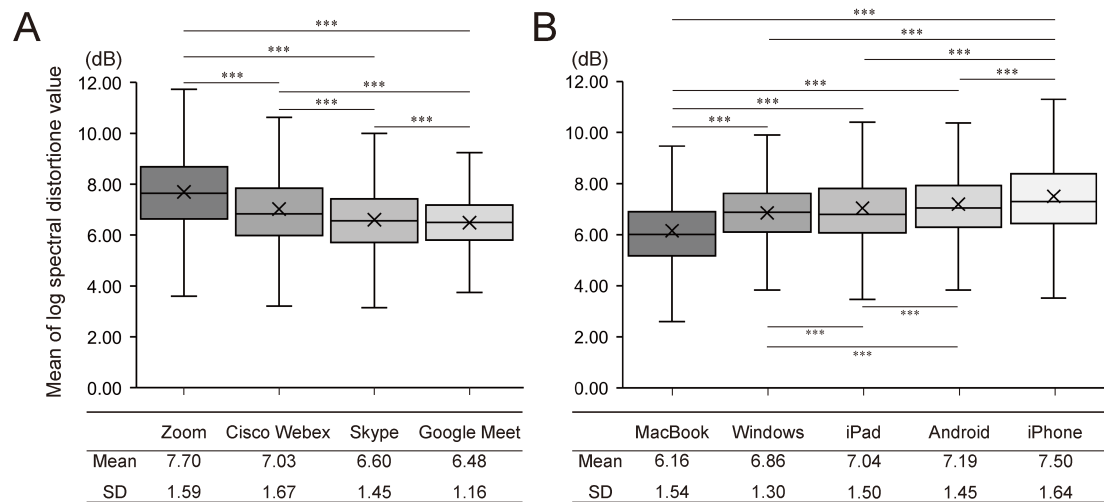


図 3 全音素の平均 LSD 値の比較

(A) 4 つの VCP 間における全音素の平均 LSD 値の比較 (n=12,925)

(B) 5 つのデバイス間における全音素の平均 LSD 値の比較 (n=10,340)

全ての VCP 間、デバイス間で有意差を認めた。(有意水準 : ***p<0.001)

図 4 は、「ふつう/φ/, /w/, /ts/, /w/, /w/」という同じ音声サンプルについて、全ての VCP とデバイスの組み合わせにおける音声波形と対数振幅スペクトルを示した例である。各組み合わせの上部が音声波形、下部が対数振幅スペクトルを示す。各組み合わせの上部の音声波形において、(A)音声サンプルと(B)Zoom®×MacBook®における入力波形を、VCP を介した他のサンプルと比較すると、低振幅の領域は小さくなるが、高振幅の領域は維持されていた。Zoom® と Android® (F)、Cisco Webex® と iPad® (J)、Skype® と Android® (P)の組み合わせでは、/φ/の冒頭で音声波形が消失し、LSD 値は 18.00 dB を超えた。各組み合わせの下部の対数振幅スペクトログラムでは、音声信号は低周波数域ではわずかな変化を示したが、高周波数域では減衰した。Zoom® と MacBook® (C) と Windows® (D)、iPhone® (G) の組み合わせでは、3800 Hz と 7000 Hz の音声信号が消失した。

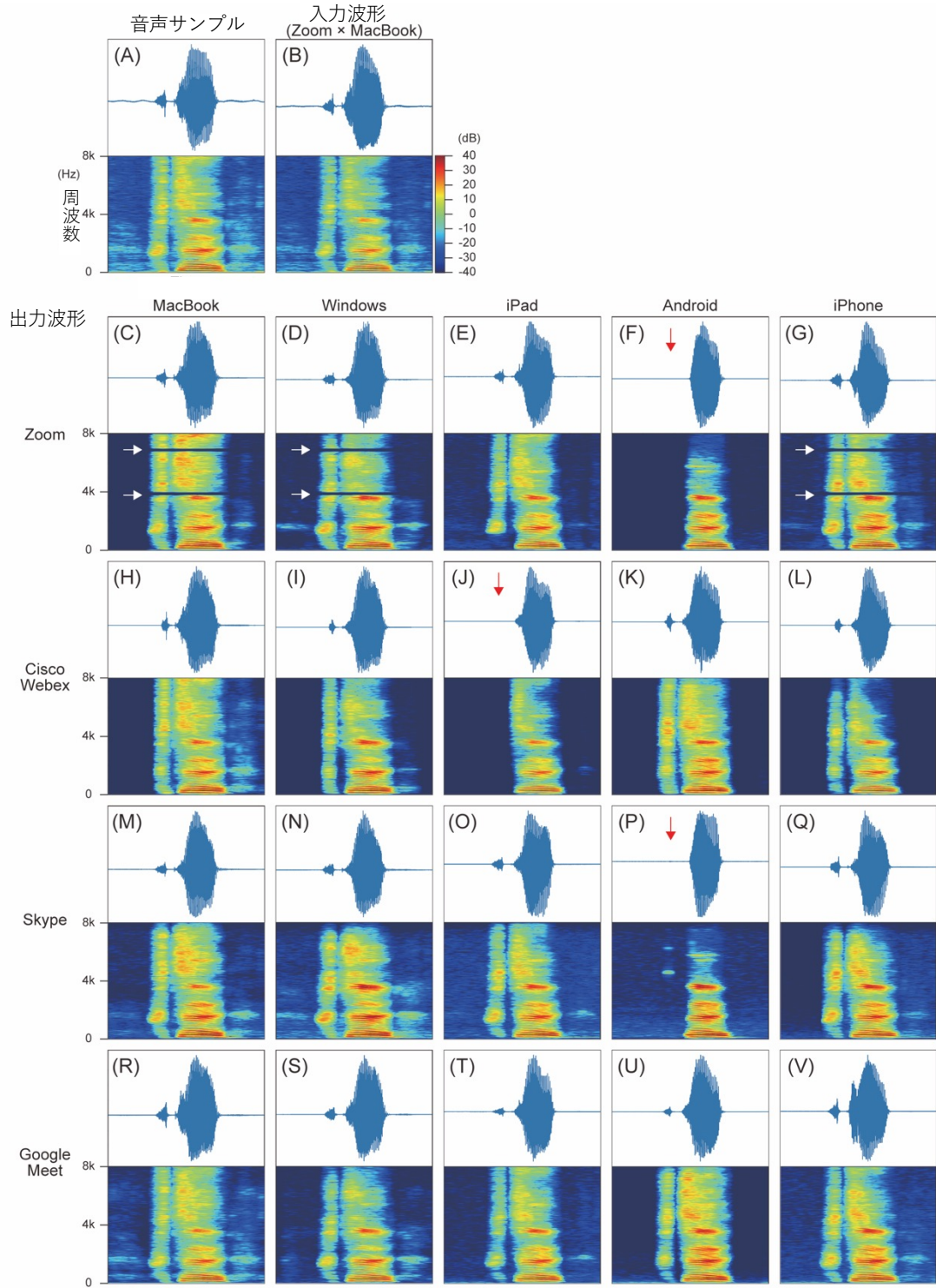


図 4 日本語の「ふつう/ɸ/, /w/, /ts/, /w/, /w」という同じ音声サンプルについて、全ての VCP とデバイスの組み合わせにおける音声波形と対数振幅スペクトルを示した例。それぞれの組み合わせにおいて、上部が音声波形、下部が対数振幅スペクトルを示す。(F)(J)(P) は低振幅領域での音声波形の消失し(↓)、(C)(D)(G)は 3800 Hz と 7000 Hz において音声信号が消失した(→)。

4.3. 各子音における伝送条件の影響

表 3 に各 VCP における子音に関する LSD の平均値(Mean)と標準偏差(SD)を示す。有声性では、全ての VCP において無声子音の平均 LSD 値は有声子音の平均 LSD 値より大きかった。図 5A に示すように t 検定において有意差を認めたが、Zoom® ($p < 0.0001$, $d = 0.376$)、Cisco Webex® ($p < 0.0001$, $d = 0.280$)、Skype® ($p < 0.0001$, $d = 0.374$)、Google Meet® ($p < 0.0001$, $d = 0.366$)であり、効果量に有意差はなかった。調音位置では、全ての VCP において歯茎口蓋音が最も高い LSD 値を示した。Zoom®, Skype®, Google Meet®では両唇音、Cisco Webex®では唇軟口蓋音が最も小さい LSD 値を示した。図 5B は Tukey's 検定の結果を示しており、全ての VCP において歯茎口蓋音が有意に大きな平均 LSD 値を示し、歯茎口蓋音-両唇音間、歯茎口蓋音-唇軟口蓋音間で効果量大であった。(サプリメンタルデータ 1)調音法では、全ての VCP において摩擦音が最も高い LSD 値を示した。Zoom®, Cisco Webex®, Skype®は破擦音、接近音、はじき音、破裂音、鼻音の順に、Google Meet®は破擦音、はじき音、接近音、破裂音、鼻音の順に LSD が大きかった。図 5B は Tukey's 検定の結果を示しており、全ての VCP において摩擦音と破擦音は有意に大きな LSD 値を示し、破擦音-鼻音間、破擦音-破裂音、破擦音-はじき音、摩擦音-鼻音間、摩擦音-破裂音間で効果量大であった。(サプリメンタルデータ 2)

表 3 各 VCP における子音の平均 LSD 値

	n	Zoom		Cisco Webex		Skype		Google Meet	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
有声性									
無声	2530	7.95	1.72	7.25	1.93	6.91	1.69	6.70	1.31
有声	3160	7.35	1.53	6.76	1.59	6.35	1.34	6.25	1.13
調音位置									
両唇音	770	6.82	1.45	6.58	1.69	6.08	1.46	6.02	1.17
歯茎音	2335	7.71	1.59	7.05	1.76	6.64	1.52	6.57	1.21
歯茎口蓋音	605	8.70	1.67	7.70	1.73	7.62	1.42	6.99	1.12
硬口蓋音	595	8.08	1.51	7.39	1.60	6.64	1.36	6.73	1.05
軟口蓋音	1075	7.28	1.46	6.80	1.76	6.34	1.56	6.29	1.29
唇軟口蓋音	185	7.19	1.39	6.49	1.30	6.21	1.12	6.06	0.94
声門音	125	7.88	1.31	7.40	1.61	6.62	1.30	6.79	1.09
調音法									
破裂音	2165	7.16	1.46	6.67	1.70	6.26	1.51	6.25	1.26
摩擦音	1015	8.95	1.50	8.09	1.85	7.55	1.52	7.28	1.07
破擦音	545	8.89	1.25	8.02	1.64	7.51	1.41	7.23	0.98
はじき音	585	7.38	1.41	6.75	1.35	6.30	1.19	6.37	1.01
鼻音	960	6.95	1.36	6.48	1.50	6.04	1.26	5.99	1.07
接近音	420	7.60	1.50	6.90	1.44	6.56	1.24	6.36	0.96

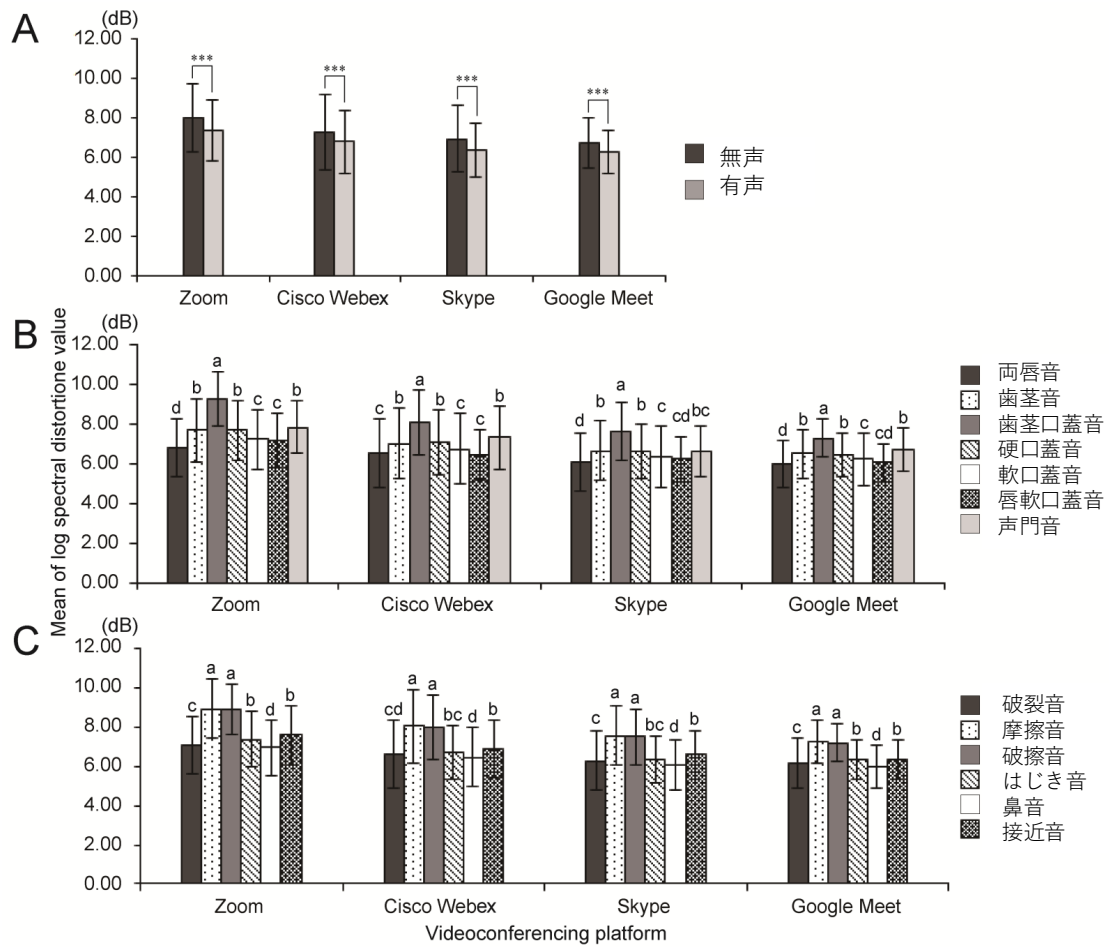


図5 4つのVCPにおける有声性(A)、調音位置(B)、調音法(C)に関する平均LSD値の比較。エラーバーは標準偏差を示す。(A) 無声子音は有声子音よりも平均LSDが大きく、t検定においてすべてのビデオ会議プラットフォームで有意差を認めた。(***) $p < 0.0001$) (B) 歯茎口蓋音は最も大きな平均LSD値を示し、両唇または唇軟口蓋音は最も小さい平均LSD値を示した。(C) 摩擦音と破擦音は最も大きな平均LSD値を示し、鼻音と破裂音は最も小さい平均LSD値を示した。エラーバー上部の異なる小文字間にはTukey's検定において有意差($p < 0.05$)を示す。

5. 小括

- 1) 全ての音素における 4 つの VCP 間の平均 LSD 値の比較では、有意に Zoom[®], Cisco Webex[®], Skype[®], Google Meet[®]の順に音声信号歪みが大きかった。Zoom[®]-Google Meet[®]間では効果量大であった。
- 2) 全ての音素における 5 つのデバイス間の平均 LSD 値の比較では、有意に iPhone[®], Android[®], iPad[®], Windows[®], MacBook[®]の順に音声信号歪みが大きかった。iPhone[®]-MacBook[®]間で効果量大であった。
- 3) 音声波形を比較すると、VCP を介した音声は低振幅の領域は小さくなるが、高振幅の領域は維持されていた。
- 4) 対数振幅スペクトログラムを比較すると、低周波数域はわずかな変化を示したが、高周波数域は減衰した。
- 5) 各子音の有声性における 4 つの VCP 間の平均 LSD 値の比較では、全ての VCP において有意に無声子音は有声子音より大きかったが、効果量に有意差はなかった。
- 6) 各子音の調音位置における 4 つの VCP 間の平均 LSD 値を比較では、全ての VCP において有意に歯茎口蓋音が最も大きく、両唇音と唇軟口蓋音が最も小さかった。歯茎口蓋音-両唇音間、歯茎口蓋音-唇軟口蓋音間で効果量大であった。
- 7) 各子音の調音法における 4 つの VCP 間の平均 LSD 値の比較では、全ての VCP において有意に破擦音と摩擦音が最も大きく、鼻音が最も小さかった。破擦音-鼻音間、破擦音-破裂音、破擦音-はじき音、摩擦音-鼻音間、摩擦音-破裂音間で効果量大であった。

研究 2

1. 目的

口蓋裂言語のうち頻度の高い PA、LA について、VCP を介した音声の音声信号歪みが言語聴覚士による構音評価に与える影響を主観的評価を用いて明らかにすることを目的とした。

2. 材料および方法

2-1. 音声データとデータ分析

インテルナ出版「口蓋裂の構音障害(Audio CD)®」に収録された PA、LA の音声サンプルを利用した。選択した音声サンプルは、PA：たちつてと(5 歳、成人)、さしすせそ(5 歳、成人)、LA：かきくけこ(7 歳、成人)、さしすせそ(7 歳、成人)の 8 つであった。これらの音声サンプルは、研究 1 で用いた ATR 503 と異なり、雑音が含まれている実臨床に近い環境下で収録されたものである。初めに、Praat®を用いて各音声サンプルの音の強さレベル(dB)を統一した。研究 1 と同様に Praat®を用いて、IPA 国際音声記号に基づくアノテーションを付与した。ここでの構音の誤りに該当する評価子音は、PA はたて/t/、つ/ts/、さすせそ/s/、LA はし/c/、き/k/であった。それぞれの音声データをオリジナルサンプルと定義した。これらのオリジナルサンプルを研究 1 と同様の VCP4 つ、デバイス 5 つを用いて、研究 1 同じ収録環境下で録音した。1 つのオリジナルサンプルに対し、対応する 20 個の評価サンプルを作成した。

2-2. 受聴試験

言語聴覚士 8 名を評価者とした。言語聴覚士の選定条件は日常的に構音訓練治療に携わっており、臨床経験 5 年以上とした。評価者は 4 施設から参加し、その内訳は施設 A から 3 名、施設 B から 2 名、施設 C から 2 名、施設 D から 1 名であった。臨床歴は 5～23 年(中央値 19 年)であった。受聴環境は、防音室または周囲雑音のない個室とした。受聴機器は各評価者保有のラップトップを使用した。ヘッドホンの周波数特性の影響を排除するため、使用機器に搭載されているスピーカーを使用するよう指定した。評価者は、初めに検証に用いない正常例の「あいうえお」を 2 回受聴し、各評価者において、音がはっきりと聞こえ、最も聞きやすいと感じられる大きさに音の強さレベルを調整した。これを各評価者の

最適な音量と定義した。オリジナル音声と、異なる VCP、デバイスを紹介した各 20 種類の評価サンプルをランダムに受聴し構音判定を行った。評価者は評価対象音声を 2 回ずつ受聴した。オリジナルサンプルと評価サンプルの回答は正常例(N)、口蓋化構音(PA)、側音化構音(LA)、声門破裂音(GS)のいずれかから選択性とした。信頼性の検証のため、同じサンプルを用い、1 週間以上空けて 2 回評価を行った。一連の判定は、機器や音量による影響を排除するため、検証途中で受聴環境、使用機器、最適な音量を変更せずに行った。疲れの影響を排除するため、20 分以上連続で評価を行わないことを条件とした。

オリジナルサンプルに対する構音判定、VCP ごとのオリジナルサンプルと評価サンプルの構音判定の一致率、評価子音の平均 LSD 値を求め、研究 1 における正常例の子音と比較した。本研究ではテレプラクティスによる影響を検証するため、構音判定の一致率は正誤に関わらず検討した。

2-3. 統計分析

SPSS(ver.30)を用いて行った。評価者間信頼性と評価者間信頼性は κ 係数を用いて比較した。VCP 間のオリジナルサンプルと評価サンプルの構音判定の一致率については、 κ 係数を用いて比較した。VCP 間における平均 LSD 値について、Tukey's 検定を用いて比較した。さらに評価子音の平均 LSD 値と研究 1 における正常例の子音の平均 LSD 値は t 検定を行い比較した。

3. 結果

3-1. オリジナルサンプルに対する構音評価

表 1 にオリジナルサンプルに対する各評価者の構音判定と正答率を示す。PA サンプルでは、評価子音/t/, /ts/は 5 歳サンプル 100%、成人サンプル 93.8%と高い正答率を示したが、評価子音/s/は 5 歳サンプル PA: 43.8%、LA: 50.0%、N: 6.3%と判定が分かれた。成人サンプルについても PA: 56.3%、LA: 31.3%、N: 12.5%と判定が分かれた。LA サンプルでは、7 歳サンプルは評価子音/k/87.5%、/c/93.8%と正答率が高かった。成人サンプルは PA や GS の判定はなかったが、評価子音/k/は LA: 31.3%、N: 68.8%、評価子音/c/は LA: 37.5%、N: 62.5%で VCP あり、正常例と判定する評価者が多かった。

3-2. 再テスト法による評価者内信頼性と評価者間信頼性

評価者内信頼性は一致率平均が 89.3%と高く、 κ 係数平均 0.83 とほぼ完全一致を示した。評価者間信頼性は、 κ 係数平均が 0.44 と中等度の一致であった。(表 2)

3-3. VCP 間の一致率の比較

表 3 に 8 人の評価者によるオリジナルサンプルと評価サンプルの構音判定の一致率と κ 係数を VCP ごとに示す。PA サンプルでは Zoom® (86.6%, $\kappa=0.70$)、Cisco Webex® (85.3%, $\kappa=0.67$)、Skype® (84.4%, $\kappa=0.65$)、Google Meet® (85.6%, $\kappa=0.68$)であり、すべての VCP でかなりの一致であった。LA サンプルでは Zoom® (87.2%, $\kappa=0.75$)、Cisco Webex® (84.1%, $\kappa=0.70$)、Skype® (84.1%, $\kappa=0.79$)、Google Meet® (86.6%, $\kappa=0.74$)であり、すべての VCP でかなりの一致であった。

3-4. 平均 LSD 値の比較

図 1 に示すように、4 つの VCP 間における平均 LSD 値は、研究 1 と同様、PA サンプル、LA サンプルともに Zoom®が最も大きい値を示し、PA サンプルでは Google Meet®、LA サンプルでは Skype®が最も小さい値を示した。Tukey's 検定の結果、PA サンプル、LA サンプルともに、Zoom®-Skype®間、Zoom®- Google Meet®間、Cisco Webex®-Skype®間、Cisco Webex®-Google Meet®間で有意差を認めた。研究 1 における正常例の平均 LSD 値と本研究の評価子音を比較すると、t 検定の結果、LA サンプル/c/を除き、PA サンプル/t/, /ts/, /s/, LA サンプル/k/は正常例と比較し有意に平均 LSD 値は大きかった。(図 2)

表 1 オリジナルサンプルに対する各評価者の構音判定

音声サンプル		施設 A				施設 B				施設 C				施設 D				正答率
		評価者 1		評価者 2		評価者 3		評価者 4		評価者 5		評価者 6		評価者 7		評価者 8		
		1 回目	2 回目	1 回目	2 回目	1 回目	2 回目	1 回目	2 回目	1 回目	2 回目	1 回目	2 回目	1 回目	2 回目	1 回目	2 回目	
PA	/t/, /ts/ 5 歳	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	100.0%
	/s/ 5 歳	LA	N	LA	LA	N	LA	PA	PA	LA	LA	N	PA	PA	PA	PA	PA	43.8%
	/t/, /ts/ 成人	PA	PA	PA	PA	PA	PA	PA	PA	LA	PA	PA	PA	PA	PA	PA	PA	93.8%
	/s/ 成人	LA	LA	LA	LA	N	N	PA	PA	PA	PA	LA	PA	PA	PA	PA	PA	56.3%
LA	/k/ 7 歳	LA	LA	PA	PA	LA	LA	LA	LA	LA	LA	LA	LA	LA	LA	LA	LA	87.5%
	/ɕ/ 7 歳	LA	LA	LA	LA	LA	LA	LA	LA	LA	LA	LA	LA	LA	PA	LA	LA	93.8%
	/k/ 成人	N	N	N	N	N	N	N	N	N	LA	LA	LA	LA	LA	N	N	31.3%
	/ɕ/ 成人	N	N	LA	LA	N	N	N	N	LA	LA	N	N	LA	LA	N	N	37.5%
正常例(N)、口蓋化構音(PA)、側音化構音(LA)、声門破裂音(GS)																		

表 2 再テスト法による評価者内・評価者間信頼性

評価者内信頼性(n=168)																		
評価者 1		評価者 2		評価者 3		評価者 4		評価者 5		評価者 6		評価者 7		評価者 8		平均		
一致率	κ	一致率	κ	一致率	κ	一致率	κ	一致率	κ	一致率	κ	一致率	κ	一致率	κ	一致率	κ	
94.0%	0.91	99.4%	0.99	80.4%	0.71	94.0%	0.90	95.2%	0.88	85.2%	0.77	68.5%	0.51	97.6%	0.96	89.3%	0.83	
評価者間信頼性(n=8)																		
1 回目		2 回目																平均
κ		κ																κ
0.44		0.44																0.44
0.21-0.4: まずまずの一致, 0.41-0.6: 中等度の一致, 0.61-0.8: かなりの一致, 0.81-1: ほぼ完全一致 完全一致																		

表 3 8 人の評価者のオリジナルサンプルと評価サンプルの構音判定の平均一致率

			1 回目		2 回目		平均	
音声サンプル	VCP	n	一致率	κ	一致率	κ	一致率	κ
PA	Zoom	160	77.5%	0.51	95.6%	0.89	86.6%	0.70
	Cisco Webex	160	75.0%	0.46	95.6%	0.89	85.3%	0.67
	Skype	160	76.3%	0.50	92.5%	0.81	84.4%	0.65
	Google Meet	160	76.3%	0.48	95.0%	0.87	85.6%	0.68
LA	Zoom	160	85.6%	0.71	88.8%	0.79	87.2%	0.75
	Cisco Webex	160	81.9%	0.65	86.3%	0.74	84.1%	0.70
	Skype	160	82.5%	0.67	85.6%	0.73	84.1%	0.70
	Google Meet	160	85.0%	0.71	88.1%	0.77	86.6%	0.74
0.21-0.4: まずまずの一致, 0.41-0.6: 中等度の一致, 0.61-0.8: かなりの一致, 0.81-1: ほぼ完全一致 完全一致								

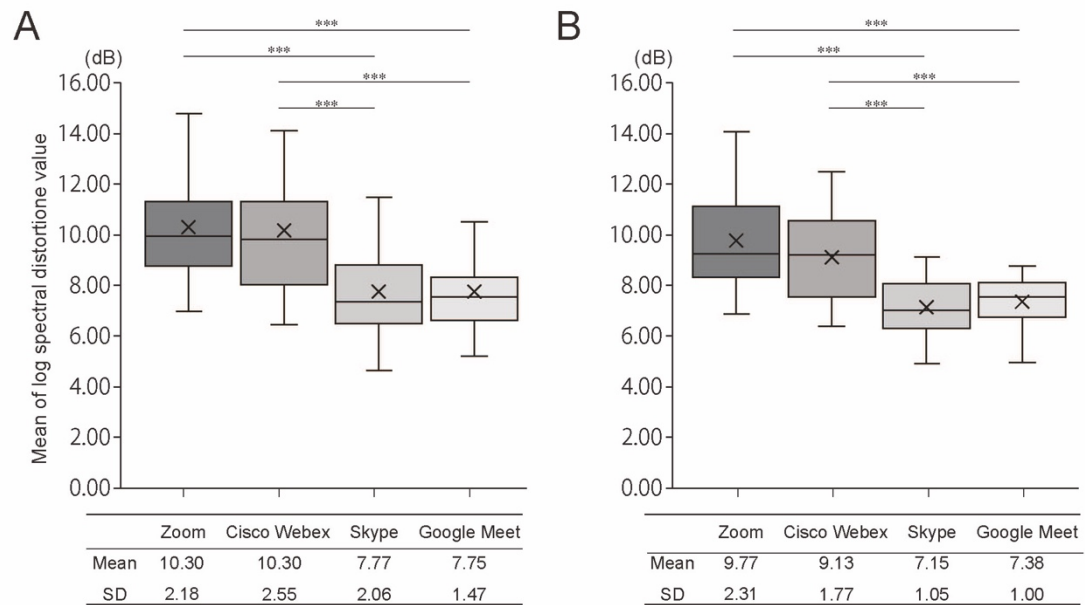


図 1 4つのVCP間における評価子音の平均LSD値の比較

(A) PA サンプル/t/, /ts/, /s/の平均LSD値 (n=280)

(B) LA サンプル/k/, /ɕ/の平均LSD値 (n=80)

有意水準*** $p < 0.0001$; Tukey's 検定

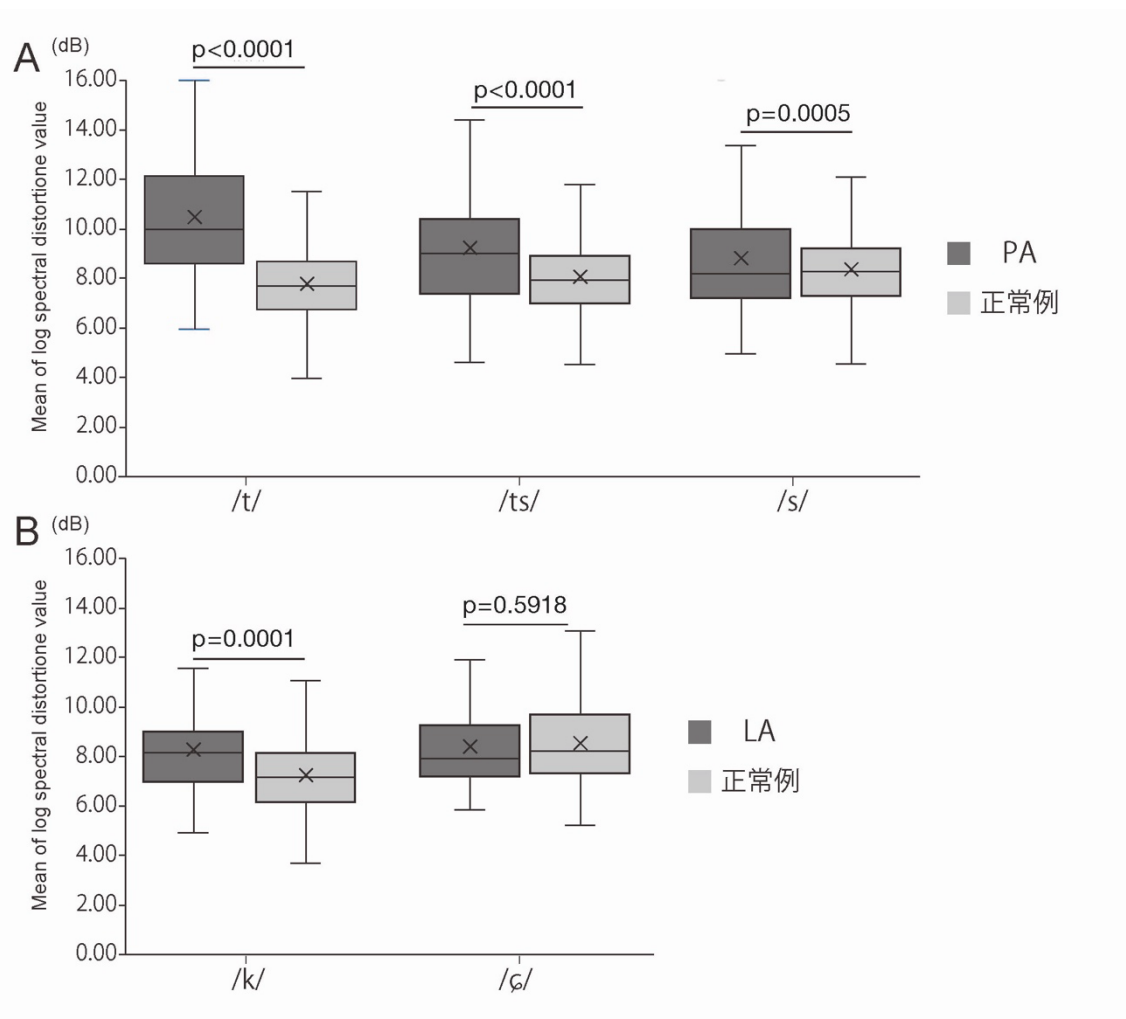


図 2 研究 1 における正常例と本研究の評価子音の平均 LSD 値の比較

(A) /t/(正常例; n=2020, PA; n=80), /ts/(正常例; n=520, PA; n=40), /s/(正常例; n=1400, PA; n=160)の平均 LSD 値

(B) k/(正常例; n=700, LA; n=40), /c/(正常例; n=960, LA; n=40)の平均 LSD 値

t 検定の結果、LA サンプル/c/を除き、PA サンプル/t/, /ts/, /s/, LA サンプル/k/は正常例と比較し有意に平均 LSD 値は大きかった。

4. 小括

- 1) PA サンプルは、評価子音/t/, /ts/において高い正答率を示したが、/s/において評価が分かれた。
- 2) LA サンプルは、成人において/k/, /c/ともに高い正答率を示したが、7 歳サンプルにおいて LA と N に評価が分かれた。
- 3) 評価者内一致率は、平均一致率 89.3% κ 係数 0.83 とほぼ完全一致を示した。
- 4) 評価者間一致率は、 κ 係数 0.44 と中等度の一致であり、評価者間である程度のばらつきがあった。
- 5) オリジナルサンプルと評価サンプルの構音判定の平均一致率は、4 つの VCP 全てで、PA、LA とともに κ 係数 0.65~0.75 と かなりの一致を示し、VCP 間で有意差はなかった。
- 6) 平均 LSD 値の比較では、PA、LA とともに Zoom[®]-Skype[®]間、Zoom[®]-Google Meet[®]間、Cisco Webex[®]-Skype[®]間、Cisco Webex[®]-Google Meet[®]間で有意差を認め、Zoom[®]と Cisco Webex[®]の音声信号歪みが大きかった。
- 7) 正常例と口蓋裂言語の平均 LSD 値の比較では、LA サンプル/c/を除き、PA サンプル/t/, /ts/, /s/, LA サンプル/k/で有意に平均 LSD 値は大きく、正常例より音声信号歪みが大きかった。

考察

研究 1 における音響学的評価では、正常例において全ての VCP、全てのデバイスで音声信号歪みが見られ、音声信号歪みの程度は VCP、デバイスによって異なることが明らかになった。更に、VCP を介することによって音声信号歪みを受けやすい音素について明らかになった。研究 2 における言語聴覚士による構音判定では、口蓋裂言語において VCP を介した音声信号歪みは生じているものの、VCP を介さない場合と構音判定はかなりの一致を示し、全ての VCP で聴覚知覚的音声評価の精度に大幅な低下を示すことはないことが明らかになった。

VCP 間の音声信号歪みに影響を与える要因

Zoom®は最も大きな平均 LSD 値を示し、音声信号歪みが大きかった。各 VCP の周波数特性は公表されていないが、本研究で検討した 4 つの VCP では周波数の減弱に違いが見られた。研究 1—図 4 で示すように Zoom®では 3800 Hz と 7000 Hz 付近で信号の消失が見られ、これらが LSD 値の上昇に影響したと考えられる。Google Meet®では最も小さな平均 LSD 値を示し、音声信号歪みが小さかった。音声強調アルゴリズム、VCP 搭載のノイズキャンセリング機能が LSD 増加に影響している可能性が考えられる [17,18]。今回の研究では、無響室内で音声データを再生および録音したにも関わらず音声信号が生じたことから、ノイズキャンセリングによる過度な抑制が、必要な音声データの信号も歪ませたことが示唆される。Google Meet®の平均 LSD 値が最も低かったのは、初期設定にノイズキャンセル機能がなかったためと考えられる。過去の研究で、インターネット帯域幅がサンプリング周波数を含む信号品質を維持する上で重要な要因であると報告しているが [20,21]、本研究ではインターネット速度を 100 Mbps 以上で一定に保ったにも関わらず音声信号歪みの程度の違いが観察されたことから、音声信号歪みの程度は各 VCP のアルゴリズムに依存することが示唆された。

デバイス間の音声信号歪みに影響を与える要因

iPhone® や Android® などのスマートフォンは、MacBook® や Windows® などのラップトップよりも大きな平均 LSD 値を示した。また、研究 1—図 4 に示すように、同じ VCP であっても、機器によって振幅スペクトログラムの周波数減衰の程度が異なっており、VCP とデバイスの組み合わせによって、音声伝送のための音声圧縮アルゴリズムが異なってい

ることが示唆された。また、今回の実験では、すべてのデバイスがスピーカーから 0.5 m の距離に設置されていることによりスマートフォンが予想以上に離れた位置にあったことから、VCP の自動音量調整機能により [30]、音量を増幅する必要があった可能性がある。

各子音の音声信号歪みに影響を与える要因

研究 1 図 4 に示すように、子音の音声信号歪みの有意差は音声波形の振幅と周波数の違いに起因することが示唆された。無声子音は有声子音と比較し、振幅が小さく、低い周波数の声帯振動を示さないため、無声子音では信号が減弱し平均 LSD 値の増加に影響したことが示唆された。また、摩擦音と破擦音は高周波数にノイズ成分を含んでおり、研究 1 図 4-(F),(J),(P)に示すように、無声両唇摩擦音/ ϕ /のような周波数の高い子音は、VCP を通すとノイズとみなされ、完全に消失する場合がある。一方で破裂音は強いバーストをもち、周波域が低い。鼻音は、更に低周波域であることから [32]、平均 LSD 値が小さかった可能性がある。

言語聴覚士による構音判定

本研究では、経験豊富で日常的に構音訓練に携わっている言語聴覚士を評価者とし、言語聴覚士が構音判定の訓練で用いる音声サンプルを用いて検証を行った。評価者内信頼性はほぼ完全一致を示したことから、個々の言語聴覚士による構音判定は一貫性があり信頼性が高いことが示唆された。しかし、評価者間信頼性は中等度の一致であり、評価者によってばらつきがあることから、聴覚知覚的な情報のみでは不十分であり、視覚的情報が必要であることや、過去の報告でもあるように評価者間の判定をさらに高い基準で一致させる必要があることが分かった [33,34]。

VCP 間の構音判定の一致率と LSD の関係

PA と LA の構音判定は、研究 2 表 3 に示すようにオリジナルサンプルと評価サンプルは全ての VCP においてかなりの一致を示し、テレプラクティスは聴覚知覚的音声評価の精度に大幅な低下を示すことはないことが明らかになった。音響学的には Zoom® と Cisco Webex® は、Skype® と Google Meet® と比較し平均 LSD 値が大きく、音声信号歪みが大きかった。しかし、構音判定の一致率に VCP 間の有意差がなかったことから、平均 LSD 7.15～10.30 dB の範囲では、どの VCP においてもテレプラクティスの有用性が示唆された。また、図 2 に示すように PA、LA は、正常例と比較し、音声信号歪みの程度が大きくなるが、

聴覚知覚的音声評価の精度には大きく影響しないことが示唆された。

臨床学的意義

過去の母音をターゲットとした研究では、テレプラクティスによって聴覚知覚的音声評価の信頼性や精度が大幅に低下することはないと報告されている [35]。本研究では、口蓋裂言語のテレプラクティスにおいて重要な子音をターゲットとして検証した。研究 1 では正常例の男性 1 名の音声データを使用した。女性や子供の声は、ピッチやフォルマント周波数が男性の声よりも高いという報告がある [36,37]。しかしながら、日本語の全子音のフォルマント周波数に男女間の有意差がないことが報告されており [24]、女性の音声データを用いた分析でも本研究と同様の結果が得られると考えられる。

口蓋裂言語の音響学的特徴は、開鼻声は声道において口腔に鼻腔が加わるため、過鼻音性を伴う母音では反共振特性やフォルマント強度の変動が生じること [38]、声門破裂音では、子音成分の消失と VOT の短縮があること [39]、口蓋化構音ではバーストの周波数が正常例と異なり、調音位置によって VOT の周波数分布にばらつきがあること [40] 鼻咽腔閉鎖機能不全を伴う口蓋裂患者では、ピッチの特徴である基本周波数が増加し、shimmer と jitter が増加することが知られている [41,42]。このような正常例とは異なる音響学的特徴を考慮すると、VCP を介した音声はノイズとして判断され、LSD 値が大きくなる可能性が考えられるため、鼻咽腔閉鎖機能不全症例での更なる検討が必要である。これまでの研究において、口蓋裂患者に対するテレプラクティスの有用性や有効性は、治療効果や臨床医、患者、保護者の満足度調査に基づいて経験的に報告されていた [14-16]。それに対して、研究 2 では口蓋裂言語 PA と LA について、VCP によってもたらされる音声信号歪みを客観指標である LSD を用いて検証し、言語聴覚士による主観評価との関係を明らかにした。

今後、段階的に信号雑音比を変化させた音声サンプルを用いた検証や、口蓋裂言語の単語や文章の音声サンプルでの検証を行い、口蓋裂患者のテレプラクティスに適切な VCP、デバイス、設定を検証する必要がある。さらに、テレプラクティスは、聴覚的な情報に加えて音声波形のリアルタイム表示などの視覚的な情報を付加することや、デジタル音声を用いた言語聴覚士の聴覚知覚判定のみに頼らない AI による構音判定の活用など、口蓋裂患者に対する言語療法の発展が期待できると考える。

結語

本研究では、LSD を用いて VCP 間・デバイス間の音声信号歪みの程度を明らかにした。PA、LA において、VCP による音声信号歪みは生じるものの、言語聴覚士による聴覚知覚的音声評価の精度には大きく影響しないことが明らかになり、テレプラクティスの有用性が示唆された。

謝辞

稿を終えるにあたり、御校閲を頂きました九州大学大学院歯学研究院口腔顎顔面病態学講座口腔顎顔面外科学分野 森山雅文 教授に深甚なる謝意を表します。また、御指導、御助言を頂きました長崎大学大学院医歯薬学総合研究科 山田朋弘 教授、九州大学大学院歯学研究院口腔顎顔面病態学講座口腔顎顔面外科学分野 大山順子 講師、九州大学病院顎顔面口腔外科 長谷川幸代 先生に深謝致します。研究に多大な御協力を頂きました九州大学芸術工学府芸術工学専攻音響設計コース 日高駿介 先生、竹久修平 先生、常に研究の協力ならびに励ましの言葉を頂きました九州大学大学院歯学研究院口腔顎顔面病態学講座口腔顎顔面外科学分野の教室員の皆様に深く感謝致します。

参考文献

- [1] Tanaka SA, Mahabir RC, Jupiter DC, Menezes JM. Updating the epidemiology of cleft lip with or without cleft palate. *Plast Reconstr Surg*. 2012; 129(3): 511e-518e.
- [2] Peterson-Falzone SJ, Trost-Cardamone JE, Karnell MP, Hardin-Jones MA. The Clinician's Guide to Treating Cleft Palate Speech. 2nd ed. Elsevier; 2016: 71-82.
- [3] Hardin-Jones M, Chapman K, Scherer NJ. Early intervention in children with cleft palate. *ASHA Lead*. 2006; 11(8): 8-32.
- [4] Broen PA, Moller KT, Carlstrom J, Doyle SS, Devers M, Keenan KM. Comparison of the hearing histories of children with and without cleft palate. *Cleft Palate Craniofac J*. 1996; 33(2): 127-133.
- [5] Kotlarek KJ, Krueger BI. Treatment of speech sound errors in cleft palate: a tutorial for speech-language pathology assistants. *Lang Speech Hear Serv Sch*. 2023; 54(1): 171-188.
- [6] Lohmander-Agerskov A, Söderpalm E, Friede H, Persson EC, Lilja J. Pre-speech in children with cleft lip and palate or cleft palate only: phonetic analysis related to morphologic and functional factors. *Cleft Palate Craniofac J*. 1994; 31(4): 271-279.
- [7] American Cleft Palate-Craniofacial Association. Parameters for evaluation and treatment of patients with cleft lip/palate or other craniofacial differences. *Cleft Palate Craniofac J*. 2018; 55(1): 137-156.
- [8] Smith AC, Thomas E, Snoswell CL, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). *J Telemed Telecare*. 2020; 26(5): 309-313.
- [9] American Speech-Language-Hearing Association. Telepractice. <https://www.asha.org/Practice-Portal/Professional-Issues/Telepractice/>.
- [10] Peng X, Xu X, Li Y, Cheng L, Zhou X, Ren B. Transmission routes of 2019-nCoV and controls in dental practice. *Int J Oral Sci*. 2020; 12(1): 1-6.
- [11] Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. *Int J Antimicrob Agents*. 2020; 55(3):105924.
- [12] Tohidast SA, Mansuri B, Bagheri R, Azimi H. Provision of speech-language pathology services for the treatment of speech and language disorders in children during the COVID-19 pandemic: problems, concerns, and solutions. *Int J Pediatr Otorhinolaryngol*. 2020; 138:110262.
- [13] Castillo-Allendes A, Contreras-Ruston F, Cantor-Cutiva LC, et al. Voice therapy in the context

- of the COVID-19 pandemic: guidelines for clinical practice. *J Voice*. 2021; 35(5): 717-727.
- [14] Bedi G, Vyas KS, Chung MT, Morrison SD, Asaad M, Mardini S. Telemedicine in international cleft care: a systematic review. *Cleft Palate Craniofac J*. 2021; 58(12): 1547-1555.
- [15] Palomares-Aguilera M, Inostroza-Allende F, Solar LR. Speech pathology telepractice intervention during the COVID-19 pandemic for Spanish-speaking children with cleft palate: a systematic review. *Int J Pediatr Otorhinolaryngol*. 2021; 144: 110700.
- [16] Hayakawa T, Imura H, Inoue C, et al. Efficacy of telepractice, an alternative therapy tool during the coronavirus disease 2019 pandemic, for speech disorders related to congenital anomalies. *Congenit Anom*. 2023; 63(6): 206-210.
- [17] Zoom Video Communications. Configuring Professional Audio Settings for Zoom Meetings. <https://support.zoom.us/hc/en-us/articles/360046244692-Background-noise-suppression>
- [18] Cisco Webex. Remove Background Noise during Webex Meetings or Webinars. <https://help.webex.com/en-us/article/n70a8os/Remove-background-noise-during-Webex-meetings-or-webinars>.
- [19] Cisco Webex. Adjust your computer mic and speaker volume in a Webex Meeting. <https://help.webex.com/en-us/article/zf86fe/Adjust-your-computer-mic-and-speaker-volume-in-a-Webex-Meeting>.
- [20] Weerathunge HR, Segina RK, Tracy L, Stepp CE. Accuracy of acoustic measures of voice via telepractice videoconferencing platforms. *J Speech Lang Hear Res*. 2021; 64(7): 2586-2599.
- [21] Tran K, Xu L, Stegmann G, Liss J, Berisha V, Utianski RL. Investigating the Impact of Speech Compression on the Acoustics of Dysarthric Speech. *Proceedings of the Annual Conference of the International Speech Communication Association, Interspeech*. 2022.
- [22] Iso K, Watanabe T, Kuwahara H. Design of a Japanese Sentence list for a speech database, *Report of the Spring Meeting the Acoustical Society of Japan*. 1988 89–90.
- [23] Itahashi S. On recent speech corpora activities in Japan. *J Acoust Soc Jpn*. 1999; 20(3): 163-169.
- [24] Nakaichi K, Wasano K. 臨床現場で用いられる単音節音源を用いた子音ホルマントの検討. *Audiology Japan*. 2021; 64(5): 386.
- [25] 厚生労働省. オンライン診療の適切な実施に関する指針. <https://www.mhlw.go.jp/content/10800000/001233212.pdf>.

- [26] 総務省. *White Paper Information and Communications in Japan*. 2022.
<https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2021/key-points.pdf>
- [27] Mondal A, Cutler R, Huang C. SureCall: Towards Glitch-Free Real-Time Audio/Video Conferencing. Section V. IEEE 18th International Workshop on Quality of Service (IWQoS). 2010.
- [28] International Phonetic Association. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge University Press; 1999.
- [29] Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans Signal Process*. 1978; 26(1): 43-49.
- [30] Wu YJ, Tokuda K. Minimum generation error training with direct log spectral distortion on LSPs for HMM-based speech synthesis. *Proceedings of the Annual Conference of the International Speech Communication Association, Interspeech*. 2008.
- [31] Bulut AE, Koishida K. Low-latency single channel speech dereverberation using U-net convolutional neural networks. *Proceedings of the Annual Conference of the International Speech Communication Association, Interspeech*. 2020
- [32] Ladefoged P, Johnson KA. *Course in Phonetics*. 7th ed. Cengage Learning; 2014: 197-226.
- [33] Otsuka N. 短音聞き書き取りによる側音化構音の評価. *Jpn J Logop Phoniatr*. 2000, 222-228.
- [34] Ogata Y, Imamura A, Fujiwara Y, et al. 口蓋化構音の Backing score による聴覚判定の試み. *J Jpn Cleft Palate Assoc*. 2023, 217-224
- [35] Dal KL, Weerathunge HR, Buckley DP, et al. Reliability and accuracy of expert auditory-perceptual evaluation of voice via telepractice platforms. *Am J Speech Lang Pathol*. 2021; 30(6): 2446-2455.
- [36] Mendoza E, Valencia N, Muñoz J, Trujillo H. Differences in voice quality between men and women: use of the long-term average spectrum (LTAS). *J Voice*. 1996; 10(1): 59-66.
- [37] Lee S, Potamianos A, Narayanan S. Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J Acoust Soc Am*. 1999; 105(3): 1455-1468.
- [38] Hirai S, Okazaki K, Arai T. A quantitative evaluation of hypernasality in children. *Jpn J Logop*. 1994; 35(2): 199-206.
- [39] Wang G, Takahashi K, Wakumoto M, Michia K. 声門破裂音の音響学特性評価. *J Cleft palate*. 1991; 16(1): 37-55.

- [40] Ozawa Y, Okazaki K. Articulatory tongue movements and acoustic characteristics of palatalized articulation in cleft palate patients. *Jpn J Logop Phoniatr.* 1994; 35(4): 322-330.
- [41] Villafuerte-Gonzalez R, Valadez-Jimenez VM, Hernandez-Lopez X, Ysunza PA. Acoustic analysis of voice in children with cleft palate and velopharyngeal insufficiency. *Int J Pediatr Otorhinolaryngol.* 2015; 79(7): 1073-1076.
- [42] Segura-Hernández M, Valadez-Jiménez VM, Ysunza PA, et al. Acoustic analysis of voice in children with cleft lip and palate following vocal rehabilitation. Preliminary report. *Int J Pediatr Otorhinolaryngol.* 2019; 126(May):109618.

サプリメンタルデータ 1 調音位置における VCP 間の Tukey's 検定と効果量

		Zoom		Cisco Webex		Skype		Google Meet	
比較		p 値	Cohen's d	p 値	Cohen's d	p 値	Cohen's d	p 値	Cohen's d
両唇音	歯茎音	<0.0001***	0.57	<0.0001***	0.26	<0.0001***	0.37	<0.0001***	0.45
	歯茎口蓋音	<0.0001***	1.75	<0.0001***	0.92	<0.0001***	1.07	<0.0001***	1.20
	硬口蓋音	<0.0001***	0.60	<0.0001***	0.31	<0.0001***	0.39	<0.0001***	0.42
	軟口蓋音	<0.0001***	0.30	0.1825	0.11	0.0050	0.17	0.0002***	0.20
	唇軟口蓋音	0.0304*	0.27	0.9960	0.05	0.9518	0.09	0.9975	0.05
	声門音	<0.0001***	0.74	<0.0001***	0.48	0.0032*	0.37	<0.0001***	0.65
歯茎音	歯茎口蓋音	<0.0001***	1.03	<0.0001***	0.62	<0.0001***	0.66	<0.0001***	0.64
	硬口蓋音	1.0000	0.004	0.0052***	0.03	1.0000	0.00	0.8409	0.06
	軟口蓋音	<0.0001***	0.29	0.0005***	0.15	<0.0001***	0.20	<0.0001***	0.24
	唇軟口蓋音	0.0003***	0.31	0.0005***	0.32	0.0023*	0.29	<0.0001***	0.41
	声門音	0.8846	0.11	0.3314	0.19	1.0000	0.01	0.4592	0.17
歯茎口蓋音	硬口蓋音	<0.0001***	1.10	<0.0001***	0.63	<0.0001***	0.71	<0.0001***	0.81
	軟口蓋音	<0.0001***	1.42	<0.0001***	0.78	<0.0001***	0.85	<0.0001***	0.89
	唇軟口蓋音	<0.0001***	1.52	<0.0001***	1.05	<0.0001***	1.05	<0.0001***	1.29
	声門音	<0.0001***	1.05	0.0003***	0.46	<0.0001***	0.72	<0.0001***	0.55
硬口蓋音	軟口蓋音	<0.0001***	0.30	0.9941	0.18	0.0014*	0.20	0.0038***	0.18
	唇軟口蓋音	0.0014***	0.35	0.0006***	0.37	0.0094*	0.31	0.0005***	0.41
	声門音	0.9290	0.11	0.6302	0.18	1.0000	0.013	0.1901	0.26

軟口蓋音	唇軟口蓋音	0.9995	0.04	0.3591	0.17	0.9213	0.09	0.3728	0.16
	声門音	0.0003***	0.42	0.0039***	0.34	0.3999	0.18	0.0001***	0.40
唇軟口蓋音	声門音	0.0026***	0.49	0.0002***	0.62	0.1889	0.35	<0.0001***	0.69

有意水準 : $p < 0.05$

Cohen's d 効果量: 0.2 =小, 0.5 =中, 0.8 =大

サプリメンタルデータ 2 調音法における VCP 間の Tukey's 検定と効果量

		Zoom		Cisco Webex		Skype		Google Meet	
比較		p 値	Cohen's d	p 値	Cohen's d	p 値	Cohen's d	p 値	Cohen's d
破裂音	摩擦音	<0.0001***	1.23	<0.0001***	0.82	<0.0001***	0.85	<0.0001***	0.86
	破擦音	<0.0001***	1.24	<0.0001***	0.82	<0.0001***	0.84	<0.0001***	0.83
	はじき音	0.0020***	0.33	0.6613	0.16	0.9890	0.21	0.0517	0.12
	鼻音	0.0340*	0.12	0.2189	0.09	0.0013	0.15	<0.0001***	0.18
	接近音	<0.0001***	0.33	0.0371*	0.16	0.0011	0.20	0.1342	0.12
摩擦音	破擦音	0.9939	0.03	0.9894	0.03	0.9925	0.03	0.9820	0.04
	はじき音	<0.0001***	1.05	<0.0001***	0.79	<0.0001***	0.89	<0.0001***	0.84
	鼻音	<0.0001***	1.37	<0.0001***	0.94	<0.0001***	1.08	<0.0001***	1.17
	接近音	<0.0001***	0.88	<0.0001***	0.68	<0.0001***	0.69	<0.0001***	0.85
破擦音	はじき音	<0.0001***	1.13	<0.0001***	0.85	<0.0001***	0.93	<0.0001***	0.85
	鼻音	<0.0001	1.45	<0.0001***	0.98	<0.0001***	1.11	<0.0001***	1.18
	接近音	<0.0001***	0.94	<0.0001***	0.72	<0.0001***	0.71	<0.0001***	0.88
はじき音	鼻音	<0.0001***	0.30	0.0330*	0.18	0.0075	0.21	<0.0001***	0.35
	接近音	0.1308	0.16	0.7470	0.10	0.0510	0.21	1.0000	0.001
鼻音	接近音	<0.0001***	0.46	0.0004***	0.27	<0.0001***	0.41	<0.0001***	0.35

有意水準 : p<0.05

Cohen's d 効果量: 0.2 =小, 0.5 =中, 0.8 =大