



原著

自動音声認識を用いた単音節明瞭度評価の検討

明崎禎輝^{1*}, 池聡², 上松智幸², 光内梨佐³, 有光一樹⁴, 平賀康嗣⁵, 金久雅史⁶¹高知リハビリテーション専門職大学 理学療法学専攻²高知リハビリテーション学院 言語療法学科³高知リハビリテーション専門職大学 言語聴覚学専攻⁴高知リハビリテーション専門職大学 作業療法学専攻⁵高知リハビリテーション学院 理学療法学科⁶高知リハビリテーション学院 作業療法学科

要旨

構音障害は、脳卒中患者や舌がん術後患者に生じる。患者の状態やリハビリテーション効果を確認するためには、適切な評価が必要である。そのため、自動音声認識による単音節明瞭度評価の使用方法を明確にすることで評価手段の選択肢を広げることになる。本研究の目的は、自動音声認識を用いた単音節明瞭度評価の信頼性と有用性を検討する。健常人 19 名（男性 7 例，女性 12 例），平均年齢 20.5 ± 1.7 歳を対象とした。評価項目は、単音節明瞭度，Oral diadochokinesis (/pa/, /ta/, /ka/, /pataka/)，発話速度を評価した。単音節明瞭度は、自動音声認識（Microsoft Stream）を用いて評価した。Pearson の積率相関係数，Spearman の順位相関係数を用い，単音節明瞭度，Oral diadochokinesis，発話速度の相関関係を分析した。有意水準は $P < 0.05$ とした。結果，Oral diadochokinesis の /pa/, /ta/, /ka/, /pataka/ はそれぞれ 6.4 ± 0.6 (5.6-7.6)， 6.6 ± 0.9 (4.2-8.0)， 5.7 ± 1.3 (2.6-7.8)， 2.1 ± 0.6 (0.6-2.8) 回/秒であった。発話速度は，平均 5.7 ± 0.6 (4.5-6.7) モーラ/秒であった。自動音声認識による単音節明瞭度評価と Oral diadochokinesis，発話速度の間では，それぞれ有意な相関関係を認めなかった。自動音声認識による単音節明瞭度評価は 100 音節中，平均 82.6(61-94) 音節を正しく認識され，20 音節が正答率 100%，23 音節が正答率 90-99%，22 音節が正答率 80-89% であった。これらのことから自動音声認識による評価は，単音節明瞭度評価として使用する上で有用であることが示唆された。

受付日 2021 年 4 月 14 日

採択日 2021 年 11 月 8 日

*責任著者

明崎禎輝

高知リハビリテーション専門職
大学 理学療法学専攻

E-mail:

akezakiteru@yahoo.co.jp

キーワード

単音節明瞭度

構音障害

自己評価

はじめに

成人における構音障害は，脳卒中による運動障害性構音障害，舌がん根治切除術・再建術後の器質性構音障害が生じることで単音節明瞭度が低下し，発話明瞭度の低下に繋がる場合がある。発話明瞭度の低下は，第三者へ音声言語で意思を伝える手段が障害され，Quality of Life の低下を引き起こす原因となる。発話明瞭度が低下している症例には，言

語聴覚士を中心とした医療従事者がリハビリテーション治療を実施することによって，明瞭度を改善させる必要がある。

がん患者は，入院中におけるリハビリテーション治療の実施は可能であるものの，外来リハビリテーションの算定が困難である。舌がん患者において，発話明瞭度が低下した状態で自宅復帰した場合には，明瞭度を改善させるための



積極的なリハビリテーション治療の介入に制限が生じる。舌がん術後患者の単音節明瞭度は¹⁾、舌部分切除症例 83.5%、舌 2/3 切除症例 59.8%、舌垂全摘症例 55.8%、舌全摘後で再建口腔底の 2 次的沈下・萎縮を防ぐことが可能であった症例 44-48%、舌全摘後で再建口腔底の 2 次的沈下・萎縮が明らかであった症例 20-37%であった。舌がん術後患者の社会復帰には、単音節明瞭度 70%以上が求められることも指摘されており²⁾、一定以上の改善が必要となる。

リハビリテーション治療効果では、舌・口底切除症例を対象に舌接触補助床を装着した上で構音訓練をすることによって単音節明瞭度の改善を得たことが報告されている³⁾。構音訓練は、構音位置付け法を用いる方法の有用性も報告されており⁴⁾、急性期では心理的負担に配慮して改善の期待できる単音節レベルから始めることが述べられている⁵⁾。練習量に関しては、自らの発話のフィードバックや自主訓練を含め十分な練習量の確保が必要である⁶⁾。構音障害に対する訓練方法は、必ずしも複雑な手技を必要とせず、鏡での視覚的フィードバックや実際に構音した音を確認しフィードバックすることで実施することも可能である。発話明瞭度が低下した症例は、入院中だけでなく、自宅復帰後にも医療従事者が提示したプログラム内容を実施することで、明瞭度を改善させる取り組みも必要となる。

単音節明瞭度の検査は、日本語無意味 100 音をアットランダムに配列した方法⁷⁾があり、患者の発話を複数名の言語聴覚士が聞き取り、正しく構音されているかを評価する。自動音声認識を利用した評価方法としては、構音障害患者に対して客観的な評価方法が検討されており⁸⁾、英語圏では自動音声認識システムの発達に伴い、頭頸部がん、喉頭切除術などによる明瞭度の低下を効果的に推定が可能となっている⁹⁻¹¹⁾。本邦ではサンプル CD を用いて文レベルに対する音声認識評価も実施されている⁹⁾。しかしながら、単音節明瞭度については検討がされておらず、自動音声認識によって適切な評価が可能であるかどうかは明らかにされていない。患者自身が自宅で単音節明瞭度の自己評価が可能であれば、自宅で患者自身による単音節レベルからのリハビリテーション治療の効果判定、現状把握に有用となる。そのため、安価で簡便であり、信頼性と妥当性の優れた単音節明瞭度評価が求められる。

本研究の目的は、自動音声認識を用いた単音節明瞭度評価が、客観的な評価ツールの一つとして用いることが可能であるか検討するため、評価の信頼性と有用性を検証する。

対象と方法

健常人 19 名 (男性 7 例, 女性 12 例), 平均年齢 20.5 ± 1.7 歳を対象とした。対象者は、全例が健常人であり、発話に障

害を有していない者であった。

評価項目は、単音節明瞭度、標準ディサースリア検査から Oral diadochokinesis と発話速度の項目を抜粋して評価した。環境設定としては、個室で評価を実施し、騒音計にて平均 44dB の雑音下で測定を行った。本研究は、高知リハビリテーション専門職大学倫理委員会の倫理基準 (R2-6) に従い実施した。なお、対象者にはインフォームドコンセントを得た後に実施した。

単音節明瞭度

自動音声認識を用いた単音節明瞭度評価を実施した。方法は、50 音+濁音・半濁音+拗音の計 100 音節を対象に音読させたものをボイスレコーダーに録音した。そして、録音した音節を Microsoft Stream の自動音声認識テクノロジーを用いて自動生成されたキャプションファイルをダウンロードし、テキスト化されたものを記録する。認識の正誤は、音読した音節が予測変換された語に含まれている、もしくは変換後の表記を評価者が音読した際に音が同一である場合に正しく認識されたと判断し、それ以外の変換を誤答とした (正答例: 同一の音; 「に」を「2」に変換、音の付加があるが同一の音も含まれる; 「みよ」を「妙」に変換、誤答例: 子音の置換; 「みよ」を「尿」に変換、子音の脱落; 「ぺ」を「え」に変換、母音の誤り; 「へ」を「日 (ひ)」へ変換)。各音節について、正しく認識された人数を正答率として算出した (正しく認識された対象数/全対象数)。

自動音声認識は認識した音節を元に、語彙、構文知識、文法知識を駆使して変換を行っている¹²⁾。そのため、単音節を読み取らせた場合に、自動的に単語やアルファベット、ローマ字などに予測変換される場合がある。

Oral diadochokinesis

対象者には、単音節 /pa/, /ta/, /ka/, /pataka/ を各 5 秒間できるだけ早く発音するように指示した。口腔機能測定装置 (健口くん; 竹井機器工業) にて総回数、1 秒間単位の平均反復回数を計測した。

発話速度

標準ディサースリア検査の「北風と太陽」を音読させ、音読開始から終了までの時間を計測し、 $(223 \text{ モーラ} + \text{繰り返し読んだモーラ数} - \text{読み飛ばしたモーラ数}) / \text{所要時間}$ の式にて発話速度を算出した。

統計解析

自動音声認識による単音節明瞭度評価と Oral diadochokinesis、自動音声認識による単音節明瞭度評価と発話速度の間における各相関関係は、Pearson の積率相関

係数, Spearman の順位相関係数を用い分析した. 統計ソフトは, IBM SPSS statistics ver 22.0 を用い, 危険率は有意水準 5%未満とした.

結果

自動音声認識による単音節明瞭度評価, Oral diadochokinesis, 発話速度は, 表 1 に示す. 単音節明瞭度評価は, 平均 82%であった.

表 1. 調査項目の結果

項目	
単音節発話明瞭度 (%)	82.6±8.1 (61-94)
pa (回数/1 秒)	6.4±0.6 (5.6-7.6)
ta (回数/1 秒)	6.6±0.9 (4.2-8.0)
ka (回数/1 秒)	5.7±1.3 (2.6-7.8)
pataka (回数/1 秒)	2.1±0.6 (0.6-2.8)
発話速度 (モーラ/秒)	5.7±0.6 (4.5-6.7)

平均 ± 標準偏差 (最小-最大)

音声認識評価において, 対象者の音節の正答率 100%は, 「え」, 「く」, 「さ」, 「け」, 「じゃ」, 「き」, 「は」, 「あ」, 「きよ」, 「よ」, 「か」, 「の」, 「ゆ」, 「が」, 「しゃ」, 「ぞ」, 「じゅ」, 「りよ」, 「す」, 「し」であった. 正答率 90-99%が「な」, 「びゅ」, 「そ」, 「に」, 「ね」, 「ふ」, 「ま」, 「や」, 「ら」, 「み」, 「しゅ」, 「ひょ」, 「と」, 「ぎ」, 「じ」, 「ず」, 「ぜ」, 「びょ」, 「で」, 「きゅ」, 「しよ」, 「せ」, 「こ」であった. 正答率 80-89%は, 「みゃ」, 「だ」, 「び」, 「きゃ」, 「ひ」, 「ちよ」, 「ひゃ」, 「じょ」, 「れ」, 「り」, 「りゅ」, 「る」, 「ろ」, 「ぎ」, 「ひゅ」, 「にゅ」, 「ぎゅ」, 「も」, 「ぼ」, 「ぬ」, 「び」, 「ぐ」であった (表 2). 音声認識評価の誤変換例は, 表 3 に示す.

自動音声認識による単音節明瞭度評価と Oral diadochokinesis, 発話速度の間では, それぞれ有意な相関関係を認めなかった (表 4).

考察

本研究では, 自動音声認識を用いた単音節明瞭度評価の信頼性と有用性について検討した. 健常人を対象に評価を実施した結果, 自動音声認識による単音節明瞭度評価は 100 音節中, 平均 82.6 音節が正しく認識された.

発話明瞭度評価において, 自動音声認識を用いた評価は信頼性が不十分であることが報告されている¹³⁾. 本研究の結果, 健常人を対象として, 本研究の正誤の判断基準にあてはめると自動音声認識による単音節明瞭度評価

は 100 音節中, 平均 82.6 音節が正しく認識され, 正しく認識された音節の幅が 61-94 音節であり, 認識の幅が認められた. このため, 今回用いた自動音声認識による単音節明瞭度評価は, 健常人を対象とした場合でも, 音節によって認識の有無に違いが生じていることが示唆された.

今回用いた自動音声認識による単音節明瞭度評価では, 「く」, 「さ」などの 20 音節が正答率 100%, 「な」, 「びゅ」などの 23 音節が正答率 90-99%, 「みゃ」, 「だ」などの 22 音節が正答率 80-89%であった. 一方, 「わ」が 10.5%, 「ちゅ」が 0%であった. 正答率が高い音節は, 疾患を有した症例に評価を実施し, 正しく認識できないと判断した場合には, 正しく構音されていない可能性が高いことが推測される. 一方, 認識率が低い音節では, 疾患を有した症例の構音が正しく認識されなくても, 正しく構音できていないとは判断できないことが示唆された. これらのことから, 今回, 自動音声認識による単音節明瞭度評価では, 正答率の高い音節と正答率の低い音節を把握し評価に使用することで, 評価の信頼性が高まることが示唆された. また単音節明瞭度は 70%以上で社会復帰が可能であることが報告されている²⁾. 本研究の結果, 正答率が約 80%以上であった音が 76/100 音節であった. そのため, 今回の評価方法は, 社会復帰を検討する上でも使用出来る可能性があるが, 言語聴覚士等による評価と自動音声認識による評価を比較していないため, 明確ではない.

糟谷ら¹⁴⁾によると, 舌切除患者の中で, 子音は切除量に関係なく /t, k, m, ʃ, r/ などが影響を受けることが指摘されている. また, 他の研究²⁾において, 舌がん患者を対象とした聴覚的明瞭度の評価では, 舌尖や舌背後部による閉鎖が得られず, 構音様式では閉鎖音が摩擦音・破擦音に, 構音点では歯茎音や軟口蓋音が両唇音・声門音などに聞き取られる傾向を示すことが報告されている. 本研究の結果, 自動音声認識による単音節明瞭度評価の正答率が約 80%以上であった音節の中には, 舌切除にて構音が困難となる音節と異聴先の音節ともに複数含まれている. また運動障害性構音障害で特に困難となる /s, ʃ, dz, ts, r/ も含まれているため, 一定の評価が可能なのではないかと考えられる.

本研究では, 自動音声認識による単音節明瞭度評価と発話速度の間で有意な相関を認めなかった. 単音節明瞭度で文章のような連続的発話を評価することは困難であり¹⁵⁾, 単音節明瞭度と発話速度は, 異なった評価であることを示していることが推測される¹⁶⁾. しかしながら, 他の研究では^{17,18)}, 発話速度を下げることで不正確な構音動作をより正確にし, 発話明瞭度を向上させる発話速

表2. 自動音声認識を用いた音節発話明瞭度評価の正答率

音節	正答率 (%)	音節	正答率 (%)	音節	正答率 (%)	音節	正答率 (%)
え	100	ふ	94.7	じょ	84.2	ぶ	78.9
く	100	ま	94.7	れ	84.2	つ	73.7
さ	100	や	94.7	り	84.2	め	73.7
け	100	ら	94.7	りゅ	84.2	ぎょ	73.7
じゃ	100	み	94.7	る	84.2	ち	73.7
き	100	しゅ	94.7	ろ	84.2	ば	73.7
は	100	ひょ	94.7	ぎ	84.2	ちゃ	73.7
あ	100	と	94.7	ひゅ	84.2	ぶ	68.4
きょ	100	ぎ	94.7	にゅ	84.2	によ	68.4
よ	100	じ	94.7	ぎゅ	84.2	む	63.2
か	100	ず	94.7	も	84.2	びゅ	63.2
の	100	ぜ	94.7	ぼ	84.2	りゃ	63.2
ゆ	100	びょ	94.7	ぬ	84.2	みゅ	57.9
が	100	で	94.7	び	84.2	へ	57.9
しゃ	100	きゅ	94.7	ぐ	84.2	ぼ	57.9
ぞ	100	しよ	94.7	みよ	78.9	い	57.9
じゅ	100	せ	94.7	た	78.9	ど	57.9
りょ	100	こ	94.7	げ	78.9	て	52.6
す	100	みゃ	89.5	ほ	78.9	びょ	52.6
し	100	だ	89.5	びゃ	78.9	ご	52.6
な	94.7	び	89.5	にゃ	78.9	お	36.8
びゅ	94.7	きゃ	89.5	ぎゃ	78.9	べ	31.6
そ	94.7	ひ	89.5	べ	78.9	びゃ	26.3
に	94.7	ちよ	89.5	う	78.9	わ	10.5
ね	94.7	ひゃ	89.5	ば	78.9	ちゅ	0

表3. 自動音声認識を用いた単音節明瞭度評価の誤変換

音節	例	音節	例	音節	例	音節	例
な	-	め	3位, 5月	と	どう	べ	B, えっ?, Pe, で, あメ
びゅ	湯	ぎょ	今日, 量, にゃー	びゅ	湯, び	きゃ	じゃ?, 逆
みよ	みゅ, 尿, 美女	や	目	ご	Gooid, こ	ぼ	ぶ, ポーン
て	で	か	100	ざ	さ	ば	あつ, Ta, バー
そ	さ	ちゅ	10, ,	ひゅ	日, Q, 湯	お	を, あつ
みゅ	三重, 自由	ぼ	と, ば, ぶ, to, 棒	じ	16	きゅ	競
た	あつ, だ	ら	が	ず	家	し	100
つ	図, 3, ず	びょ	秒, 表, 今日	ぜ	じ	しよ	Hi
に	Mi	り	Bi, Dg, Rei	う	ぶ, え, -	ひ	じゃあ, 去
げ	K	ち	G, -	びょ	よ, 美容	ちゃ	じゃあ, じゃ, -
ね	-	りゅ	湯	だ	Ta	ぶ	ふう, やで, ベ
じょ	じゃあ, 超	る	~, 琉	じゅ	100	ぬ	ね
ふ	う	み	自由	で	-	ちよ	所, 場
へ	ふえつ, えつち, 日	しゅ	,	にゅ	湯, 二	び	B
ほ	おう, 4, を	ろ	No, う	ど	To, Dawn	わ	は?, あ
びゃ	Via, きゃ, や, いや	い	Sh,	ば	あつ, ま, は?	ひゃ	いや, シャア
ま	な	みゃ	やあ, にゃ	ぎゅ	湯, Q, ゲー	ぐ	9, Q, ゲ
れ	で, えつ	ぎゃ	いや, や, ー, じゃ?	び	右	せ	シェ
びゃ	やあ, いや, や, 部屋, び	べ	ペイ, テイ, え?	ぶ	Akb, る?, 米	りゃ	や, ギャ, やあ, いや
む	ぬ, ミュー, 目	ひょ	今日	によ	にゃん, 妙, 日本, にゃ	こ	ゆえ
にゃ	いや, やあ	ぎ	好き, に	も	ま, の		

-: 認識困難

表4. 自動音声認識による単音節明瞭度評価とOral diadochokinesis, 発話速度の関係

	pa	ta	ka	pataka	発話速度
単音節明瞭度	$r=0.403^a$ $p=0.087$	$r=0.063^a$ $p=0.797$	$r=0.106^a$ $p=0.667$	$r=0.139^b$ $p=0.571$	$r=0.011^a$ $p=0.965$

^a Pearsonの積率相関係数

^b Spearmanの順位相関係数

度の調整法の効果を示している。日本語は音節文字であり、拗音などを除いて1音節が1文字とほぼ対応しており、1音節ごとに構音されるという言語特性から特に有効であったとしている。そのため、単音節明瞭度はこれらの代償手段を導入する際には特に重要となり、評価および訓練は重要であるといえる。疾患を有する患者に対して、単音節明瞭度を高めるリハビリテーション治療を実施する場合に、実際の会話を想定した上で、発話速度の調整法を併用する代償手段の獲得にも繋がることが考えられた。

本研究では、発話の生成に必要な運動機能を生理学的に測定する評価項目として、自動音声認識を用いた単音節明瞭度評価とOral diadochokinesis, 発話能力である発話速度を評価し、単音節明瞭度評価との関連性を検討した。しかしながら、自動音声認識を用いた単音節明瞭度評価とOral diadochokinesis, 発話速度との間では有意な相関関係を認めなかった。本研究の対象者は、健常成人であったことが結果に影響したことが考えられる。またOral diadochokinesisは構音器官の運動速度と規則性を評価するものであるため¹⁹⁾、単音節明瞭度検査とは異なる側面を評価していることが原因であると推測される。

今回、自動音声認識による単音節明瞭度評価は、100音節のうち、認識されやすい音節が認められた。自動音声認識は、全ての音節を正しく認識することが困難であるものの、正しく認識が可能な音節は、評価として用いることが可能である。自動音声認識による単音節明瞭度の自己評価は、疾患を有する症例が、自分自身で実施したリハビリテーション治療効果の検証、および現状把握として評価する上で有用であることが示唆された。

最後に本研究の限界点について述べる。第一に、単音節明瞭度評価で用いた自動音声認識はMicrosoft Streamを使用しており、他の自動音声認識や言語聴覚士が評価した場合の比較は実施していない。第二に、対象者は健常人であり、脳卒中やがん疾患を有する症例を対象としていない。第三に、健常人を対象に評価を実施したが、

正しく認識されていない音節には、構音様式・構音点ともに共通点がなく、「お」を「を」と認識するなど、音が近似する誤認識、「ちゅ」を「,」と音節ではないものに誤認識する単音節が認められ評価の限界がある。今後、更に検討する必要がある。

利益相反

開示すべき利益相反はない

謝 辞

本研究に協力して下さった対象者に感謝いたします。また本研究は、高知リハビリテーション専門職大学「共同研究助成金」の助成を受け実施した研究成果の一部である。

まとめ

本研究では、自動音声認識による単音節明瞭度評価を実施した。対象者は、各音節で正答率が異なる結果であった。疾患を有する症例に用いる場合には、正答率の高い音節は、適切に評価されている可能性が高いことが考えられた。自動音声認識による単音節明瞭度評価は、疾患を有する症例が自己で評価する上で実施可能な評価方法であることが推測された。

文献

- 1) 今野昭義, 花沢秀, 吉野泰弘・他: 舌切除後の舌・口腔底再建術と術後の構音機能および咀嚼機能の評価. 耳鼻 34: 1393-1407, 1988
- 2) 熊倉勇美: 舌切除後の構音機能に関する研究: 舌癌 60 症例の検討. 音声言語医学 26: 224-235, 1985
- 3) 今井智子, 山下夕香里: 舌・口底切除症例に対する舌接触補助床の有効性. 聴能言語学研究 9: 1-9, 1992
- 4) 今井智子: 口腔・中咽頭腫瘍 (4) 治療・訓練 (2) (小野寺富子・監: 言語聴覚療法臨床マニュアル, 改訂第2版). 協同医書出版社, 2011



- 5) 澤美菜子, 神田享: 第8章構音障害 2. 器質性構音障害(口腔・中咽頭がん) 訓練 (2) 言語聴覚療法臨床マニュアル, 改訂第4版. 協同医書出版社, 2018
- 6) 外山慶一: 第8章構音障害 3. 運動障害性構音障害, 訓練 (3) 言語聴覚療法臨床マニュアル, 改訂第4版. 協同医書出版社, 2018
- 7) 降矢宜成: 言語障害の語音発語明瞭度(語明度)に関する研究. 日耳鼻 61: 1923-1948, 1958
- 8) Tu M, Wisler A, Berisha V, et al.: The relationship between perceptual disturbances in dysarthric speech and automatic speech recognition performance. J Acoust Soc Am 140: EL416, 2016
- 9) Stelzle F, Maier A, Nöth E, et al.: Automatic quantification of speech intelligibility in patients after treatment for oral squamous cell carcinoma. J Oral Maxillofac Surg 69: 1493-1500, 2011
- 10) Schuster M, Haderlein T, Nöth E, et al.: Intelligibility of laryngectomees' substitute speech: automatic speech recognition and subjective rating. Eur Arch Otorhinolaryngol 263: 188-193, 2006
- 11) Windrich M, Maier A, Kohler R, et al.: Automatic quantification of speech intelligibility of adults with oral squamous cell carcinoma. Folia Phoniatr Logop 60: 151-156, 2008
- 12) 山口優実, 荒川友美, 川口美奈子・他: 音声認識ソフトウェアを利用した客観的発話明瞭度評価. 耳鼻と臨床 64: 121-129, 2018
- 13) 山口信, 古賀由紀: HMM 音声認識ソフトを構音障害評価に使用する可能性. 柳川リハビリテーション学院・福岡国際医療福祉学院紀要 2: 16-19, 2006
- 14) 糟谷政代, 岡本善博, 金田敏郎・他: 舌癌摘出症例の構音障害(第2報) 舌の切除量が「ことば」に及ぼす影響. 日本口腔科学会雑誌 29: 398-406, 1980
- 15) 岡本途也, 降矢宜成, 小倉孝: 語音発語明瞭度検査法について. 耳鼻咽喉科 28: 587-591, 1956
- 16) 西尾正輝, 新美成二: Dysarthria における発話明瞭度に関する検討. 音声言語医学 42: 9-16, 2001
- 17) 西尾正輝, 田中康博, 阿部尚子・他: Dysarthria の言語治療成績. 音声言語医学 48: 215-224, 2007
- 18) 西尾正輝: ディサースリアの基礎と臨床: 第3巻臨床実用編, インテルナ出版, 東京, 2006
- 19) Duffy JR: Motor speech disorders. Mosby, 1995



Original article

Examination of monosyllabic intelligibility evaluation using an automatic speech recognition

Yoshiteru Akezaki^{1,*†}, Satoshi Ike², Tomoyuki Uematsu², Risa Mitsuuchi³, Kazuki Arimitsu⁴, Hiraga Yasushi⁵, Masashi Kanehisa⁶

¹ Division of physical therapy, Kochi professional university of rehabilitation

² Department of speech, language and hearing pathology, Kochi rehabilitation institute

³ Division of speech, language and hearing pathology, Kochi professional university of rehabilitation

⁴ Division of occupational therapy, Kochi professional university of rehabilitation

⁵ Department of physical therapy, Kochi rehabilitation institute

⁶ Department of occupational therapy, Kochi rehabilitation institute

ABSTRACT

Speech disorders occur in patients with stroke and tongue cancer surgery. Appropriate evaluation is required to confirm the patient's condition and effect of rehabilitation. Clarifying the use of monosyllabic intelligibility by automatic speech recognition, therefore, will expand the choice of evaluation tools. The purpose of this study was to examine the reliability and usefulness of monosyllabic intelligibility evaluation using an automatic speech recognition. The participants comprised 19 healthy individuals (7 males and 12 females), with an average age of 20.5 ± 1.7 years. The items evaluated were monosyllabic intelligibility, oral diadochokinesis (/pa/, /ta/, /ka/, /pataka/), and speaking rate. Monosyllabic intelligibility was evaluated using an automatic speech recognition (Microsoft Stream). Pearson product-moment correlation coefficient and Spearman's rank correlation coefficient were performed to examine the correlations of monosyllabic intelligibility, oral diadochokinesis, and speaking rate. The level of significance was set at $P < 0.05$. As a result, /pa/, /ta/, /ka/, and /pataka/ of oral diadochokinesis were 6.4 ± 0.6 (5.6-7.6), 6.6 ± 0.9 (4.2-8.0), 5.7 ± 1.3 (2.6-7.8), and 2.1 ± 0.6 (0.6-2.8) time/ one second, respectively. Speaking rate was 5.7 ± 0.6 (4.5-6.7) mora/ one second. pataka was 2.1 ± 0.6 (0.6-2.8) time/ one second. Monosyllabic intelligibility was not correlated with oral diadochokinesis and speaking rate. The evaluation of monosyllabic intelligibility by the software correctly recognized an average 82.6 ± 8.1 (61-94) out of 100 syllables. Percentage of correctly recognized syllables by the automatic speech recognition was 100% for 20 syllables, 90-99% for 23 syllables, and 80-89% for 22 syllables, respectively. Therefore, the results suggest that evaluation by automatic speech recognition is useful for the evaluation of monosyllabic intelligibility.

Key words: monosyllabic intelligibility, speech disorders, self-assessment