

モーションキャプチャによる 日本手話の手の位置の音素に関する分析

西牧樹生¹ 堀内靖雄¹ 原大介² 黒岩眞吾¹

概要: 手話はろう者が用いる自然言語であり音声言語を参考に研究が行われている。言語の特徴の一つである二重分節を手話に当てはめると、手話の音素は「手型」、「手の位置」、「手の動き」の3種類が存在すると考えられている。本研究では言語学的に定義される手話の音素が3次元空間でどのように表現されるかを解明することを目的としており、その第一歩として身体部位を基に定義される位置の音素の分析を行った。分析に使用するデータとしてモーションキャプチャで撮影された6,359の手話単語が収録されているデータベースと手話の音素表記が記述されている辞書データを用いる。位置の音素に対応する手の位置座標を抽出するため、各手話音節に対して、音素位置を表現している手の代表部位を手型に基づいて決定した。動きのある手話音節については、接触や身体との接近に基づいて、音素位置座標を抽出するタイミングを決定した。1,295音節のデータを分析した結果、言語学的に設定された位置の音素が3次元空間でも対応する身体部位の周辺で表現されていることが明らかとなった。

Analysis of Hand Position Phonemes in Japanese Sign Language Using 3D Motion Capture

NISHIMAKI TATSUKI^{†1} HORIUCHI YASUO^{†1}
HARA DAISUKE^{†2} KUROIWA SHINGO^{†1}

1. はじめに

手話は主にろう者が使用する自然言語である。手話の言語学的な研究はStokoeの研究を機に、以降の研究では音声言語の研究を基にして手話の解析が行われてきた。従来の言語学において、文は意味を持つ単位である語や形態素から構成され、語や形態素は意味を持たない最小単位である音素から構成されるという二重分節の構成が言語の特徴の一つであると考えられている。このような言語学的な構成を手話に当てはめると、手話では「手型」、「(手の)位置」、「(手の)動き」の3種類が音素であると考えられている。

著者の一人は手話単語の音素を記述した辞書データとして「日本語—手話辞典コーディング[1]」を作成している。このデータでは「日本語—手話辞典[2]」に記載される手話単語について、ろう者の認識に基づいて音素の記述が行われている。

一方、著者らは科学研究費基盤研究(S)「多用途型日本手話言語データベース構築に関する研究[3]」において、モーションキャプチャを用いた高精度な手話の3次元動作のデータベースの作成に携わってきた。

本研究では言語学に基づき定義が行われた手話の音素と実際にろう者が行った手話動作から得られる物理量の関係を明らかにすることを目的としている。その第一歩として、本研究は身体部位を基に定義される位置の音素が3次元空間上でどのように表現されるか分析することを目的とする。

2. 使用するデータ

2.1 3次元動作データ

本研究では科学研究費基盤研究(S)「多用途型日本手話言語データベース構築に関する研究[3]」において、高精度な光学式モーションキャプチャで計測が行われた日本手話の単語と対話のデータベースから単語データを用いる。単語データは、男性と女性の各1名の手話母語者によって表出される手話動作を東映トークン研究所にて1,600万画素のカメラ42台を用いて2×2×2m³の領域を空間分解能0.5mm、時間分解能119.88fpsの精度で計測されている。日本語の語彙を基に設定された4,873ラベルに対応する6,359の手話動作が収録されている。身体には112個の再帰性反射マーカーが取り付けられており、各マーカーの3次元座標が時系列的に記録されている。

2.2 辞書データ

本研究では辞書データとして「日本語—手話辞典コーディング[1]」を用いる。このデータは「日本語—手話辞典[2]」に登録されている全ての手話単語を音素レベルで記号化することを目指している。1単語が複数音節から構成される場合があるため、音節を基本の単位としている。例えば(できる)は右手指先で左胸と右胸へ順に触れる動作であるが、2回の接触の間の動きを「わたり」として左胸に触れる「できる(前半)」と右胸に触れる「できる(後半)」の2音節として登録されている。ただ日本手話の大半は1音節で構成されるため単語と音節は特に区別はしない。音節は全部で約2,600音節が登録されている。また音節に対して音素レベルで記号化を行うために「手型」、「位置」、「動き」や

¹ 千葉大学
Chiba University
² 豊田工業大学
Toyota Technological Institute

「掌・第3中手骨の向き」, 「両手・片手の区別」, 「接触」などのカテゴリが定義されている。

続いて本研究で用いる「位置」について記述する。位置の音素は表1に示すように身体部位を基にした20種類の音素と身体の正面の空間を示すニュートラルスペース2種類の音素が定義されている。各音節には位置の音素が1個割り当てられている。割り当てる位置の音素の判断は手を身体から離して意味が通じなくなる音節であるかを基準に行い、意味が通じなくなる場合は身体部位を基にした20種類の音素の中から選ばれて、意味が通じる場合はニュートラルスペースとして手の高さを区別した2種類のどちらかが選ばれる。

表1 位置の音素

音素	身体部位
fc	顔全体・顔の周り
uf(dk)	顔の上半分 (おでこから頭頂部)
uf(kmk)	顔の左または右上半分 (こめかみ)
uf(m)	顔の上半分 (目)
mf	鼻 (鼻下)
lf(h)	顔の下半分 (頬)
lf(k)	顔の下半分 (口・唇)
lf(g)	顔の下半分 (あご・あごの下)
er	耳・耳周辺
nk	首
tk(u)	胴体 (肩から胸辺りまで)
tk	胴体 (胸から下)
lg	足 (おもに太もも)
ua	上腕 (肩以外の腕から肘の上まで)
el	肘
la	下腕 (肘から手首手前まで)
am	腕全体 (上腕下腕にまたがる場合)
wr	手首内側・手首外側 以外すべての手首
iw	手首内側 (脈を取る方)
ow	手首外側 (手の甲側)
ns(u)	ニュートラルスペース上部 (顔の上部より上)
ns	ニュートラルスペース

2.3 分析対象の手話動作

本研究では多用途型日本手話言語データベースに収録されている6,359の手話動作の中で日本語-手話辞典コーディングと対応付け可能な音節を分析に用いる。各音節には22種類の位置の音素から1個が割り当てられており、本研究では頭部と胴体に関する13種類の音素 (fc, uf(dk), uf(kmk), uf(m), mf, lf(h), lf(k), lf(g), er, nk, tk(u), tk, lg) を対象として分析を行った。なお「lg」を持つ音節は

計測された動作データには存在しなかった。これら以外の音素のうち、非利き腕に関する音素 (ua, el, la, am, wr, iw, ow) は非利き腕の構え方によって手の位置座標が変化するため、対象外とした。それに加え、ニュートラルスペース (ns(u), ns) は身体の前面で自由に動くことが可能であることが予測されるため、対象外とした。また、非手指動作を持つ音節は手の位置座標の抽出のために手以外の部位の動きを扱う必要があるため対象外とした。そのため、本研究では表2に示す1,295音節を分析対象とする。ただし、一部の音節には男性か女性的一方のみで表出が行われたものが存在する。

表2 音節の数の詳細

分類	音節数
対応付け可能な音節	5,972
ニュートラルスペース	4,552
非利き腕に関する音素	115
非手指動作	10
分析対象	1,295

3. 手の位置の音素の分析

3.1 座標を抽出する代表部位の設定

先行研究[4]の結果、手の中心の座標は音素の身体部位との関連性が低いことが分かっている。そのため音節ごとに位置の表現において重要と考えられる手の接触部位を主に用いて分析を行う。手の接触部位の判断は辞書データの「利き手の接触する部位」の利用する。ここでは手の接触部位を「指先」、「指の掌側」、「指の手の甲側」、「指の側面」、「掌側 (指は含んでも含まなくても可)」、「手の甲側 (指は含んでも含まなくても可)」、「手の不明な部位」や「手首」などの腕の部位に分類しており、指に関する接触は5指のどの指を用いたかの詳細も記載されている。なお、モーションキャプチャにおいてマーカーは手の甲側に付けられているため掌側、手の甲側、側面については区別を行わない。「手の不明な部位」については、本研究では伸ばした指の指先を用いる。また、接触しない音節については「利き手の接触の部位」が使用できないため、その音節に関しては5指を握る手型の場合は手の中心の座標、指を伸ばす手型の場合は伸ばした指の指先を用いる。ただし、「朝飯前」や「見る②」、「ねずみ」、「宣伝」のような指先より掌や手の甲が身体に近い音節は手の中心の座標を用い、「ビール②」は親指と小指を伸ばすが、親指のみが口に接近するため親指の指先を用いる。手の接触部位と座標を取り出すマーカーの対応関係を表3に示す。

表 3 手の接触部位とマーカーの対応

手の接触部位	座標決定に用いるマーカー
指 1 本の指先	該当の指の指先マーカー 1 個
指 2 本の指先	該当の指の指先マーカー 2 個 の中点
親指, 人差指, 中指の指先	親指, 人差指, 中指の指先マーカー 3 個の重心
親指除く 4 指の指先	中指, 薬指の指先マーカー 2 個の中点
小指除く 4 指の指先	親指, 中指の指先マーカー 2 個の中点
5 指の指先	親指, 中指, 薬指の指先マーカー 3 個の重心
親指 1 本の全体	親指第 1 関節のマーカー 2 個の中点
親指除く 指 1 本の全体	該当の指の第 1, 2 関節のマーカー 2 個の中点
親指, 人差指の全体	親指, 人差指の指先と付け根関節のマーカー 4 個の重心
親指除く 2 指の全体	該当の指の第 1, 2 関節のマーカー 4 個の重心
人差指, 中指, 薬指の全体	中指の第 1, 2 関節のマーカー 2 個の中点
親指除く 4 指の全体	中指, 薬指の第 1, 2 関節のマーカー 4 個の重心
5 指全体	中指の第 1, 2 関節のマーカー 2 個の中点
付け根関節を曲げた 5 指の全体	親指の第 1 関節, 中指の第 1, 2 関節のマーカー 4 個の重心
手の全体	手の甲のマーカー 2 個の中点
手首	手首のマーカー 2 個の中点

3.2 接近タイミングの設定

動きを伴う手話音節では位置に併せて動きも同時に表現されることから, 手の座標は手話動作中に変化する. そのため動きのどの時点での手の位置座標を抽出するかを辞書データの「接触」を基にした接近タイミングを用いて分析を行う. ただし, 一部の音節では接触しない場合があり「接触」が利用できないため, その音節に関しては動きの種類や方向を基に設定を行った. 表 4 に接近タイミングの詳細を示す. そして, 設定した接近タイミングを基に 3 次元動作データより代表の部位の速さの変曲点となる時点を探して座標の抽出を行う. なお, 「動きの全体」と設定した音節は始点から終点までの動きの軌跡の重心を手の位置座標として用いる. 「動きの真中」と設定した音節は表 5 で設定する身体の部位の代表的な座標との距離が最短となる時点における手の座標を用いる.

表 4 接近タイミングの設定

状態	動きと接触の様子	接近タイミング
接触あり	x	始点
	x	終点
	x x	始点と終点
	xxx	動きの全体
	x	動きの真中
接触なし	前方向の動き	始点
	提示・押印動作	終点
	後ろ方向の動き	
	往復の動き	動きの全体
	円の動き	
	微動の動き	
	曲がる動き	
	他の動き	

表 5 身体の部位の代表的な座標の設定

身体の部位	座標決定に用いるマーカー
uf(dk)	左右の眉頭のマーカー 2 個の中点
lf(k)	口の周囲のマーカー 8 個の重心
tk(u)	両肩, 腹のマーカー 3 個の重心
tk	腹のマーカー 1 個

3.3 結果と考察

位置の音素として, 手の 3 次元座標に関して分析を行うことも可能であるが, 手話対話は対面して行われることから位置の音素は上下左右平面上における位置が重要であると考え, 本研究では上下左右の 2 次元平面上で分析を行う. また姿勢の変化による座標値への影響は小さいと仮定して, 鼻先のマーカーを原点とした相対的な位置を扱うこととする. 各位置の音素の手の代表部位の座標の散布図を図 1~図 8 に示す. なお身体部位のおおよその位置を示すため, 頭部は帽子の裾, 耳, 咬筋, あご先にあたる部分に存在するマーカーによって, また, 胴体は襟首, 肩, 腰のマーカーによって示している. 目, 口はその外周のマーカーで位置を表示している.

音素ごとに対応する身体部位の周辺に分布していることから位置の音素が手の 3 次元座標を基に区別が可能であることが分かる.

しかし, 一部の音節は身体部位から離れる事象が確認される. また身体上で広範囲に及ぶ部位である「fc」, 「tk(u)」や「tk」の音素は分布の様子が異なる. それらの事項を個別に考察を行う.

三角形のマーカーが表す音節 (図 1(a)と図 2(a)「uf(dk)」の〈アフリカ〉・〈失敗②〉・〈若い〉, 図 3(a)と図 4(a)「lf(h)」

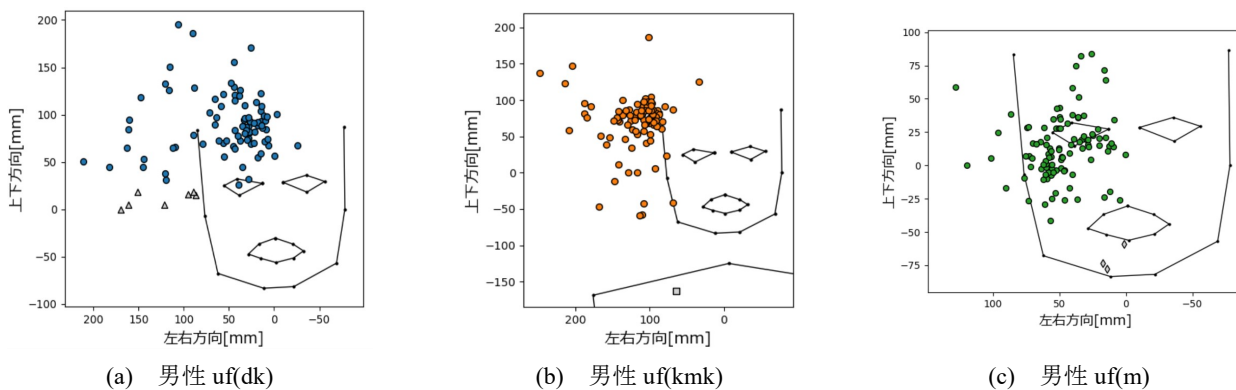


図 1 男性の顔上半分の部位に関する分布

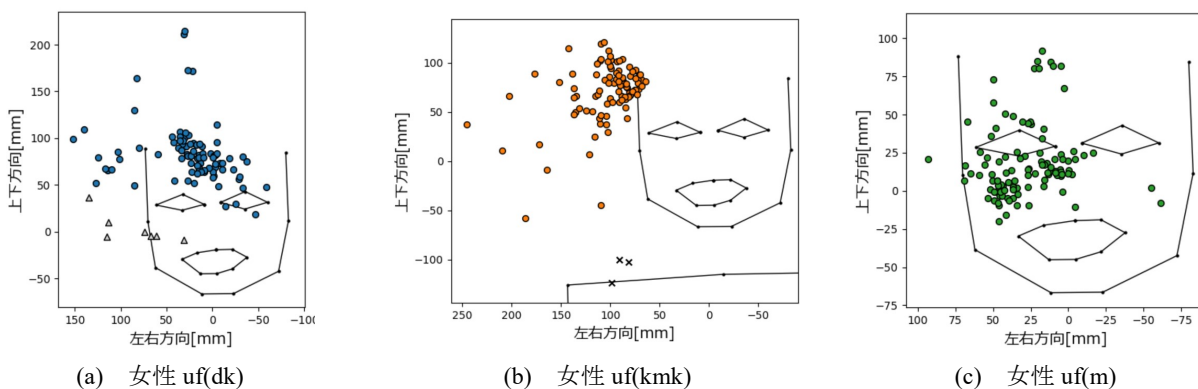


図 2 女性の顔上半分の部位に関する分布

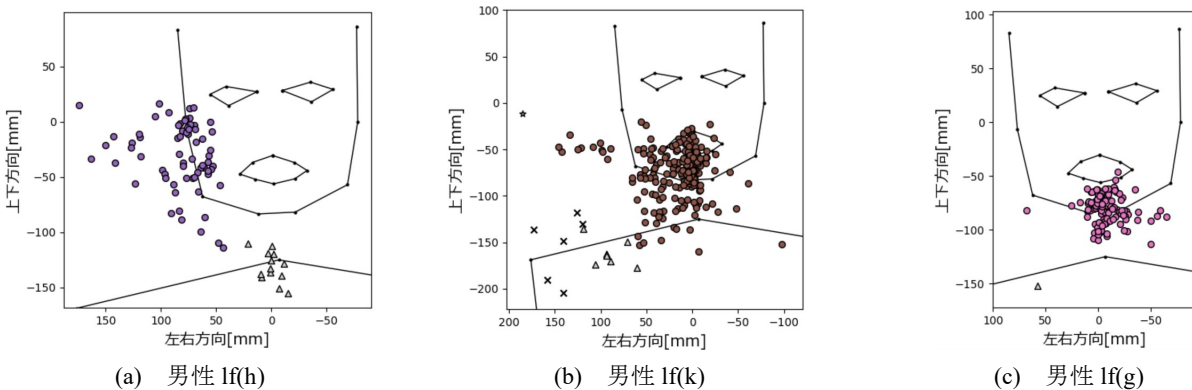


図 3 男性の顔下半分の部位に関する分布

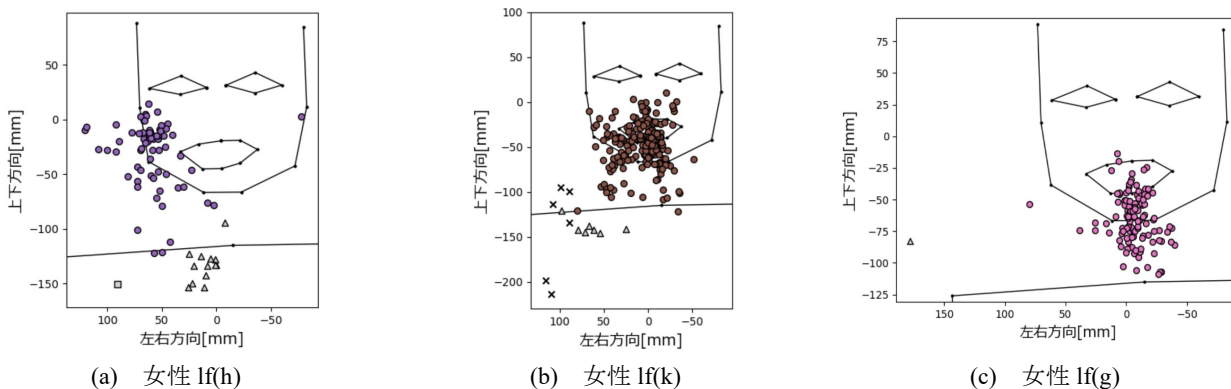
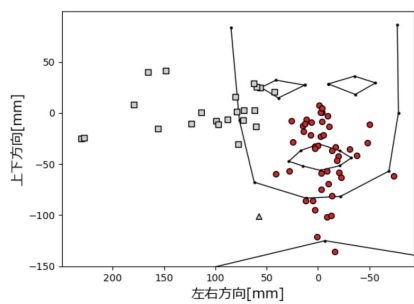
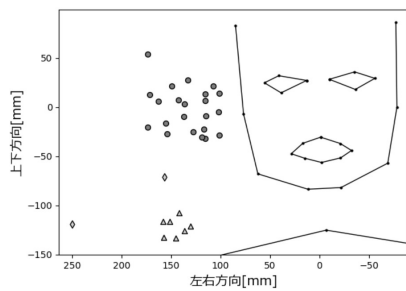


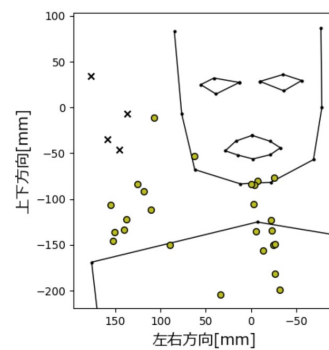
図 4 女性の顔下半分の部位に関する分布



(a) 男性 mf

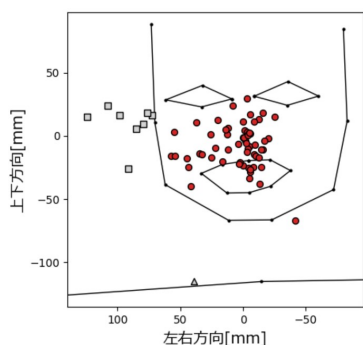


(b) 男性 er

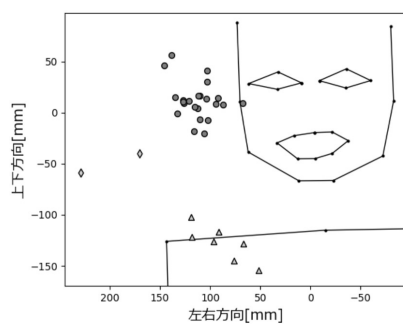


(c) 男性 nk

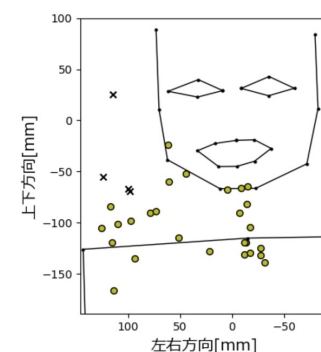
図 5 男性の顔の他の部位に関する分布



(a) 女性 mf

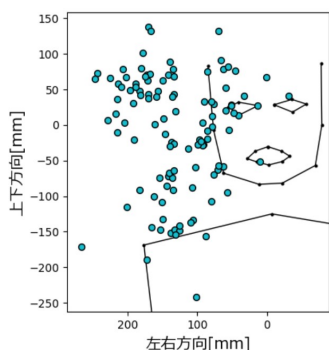


(b) 女性 er

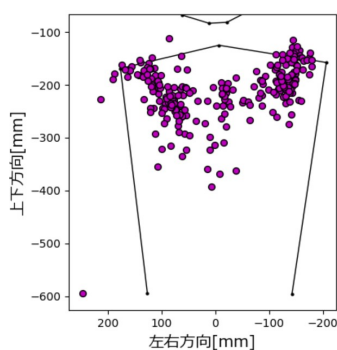


(c) 女性 nk

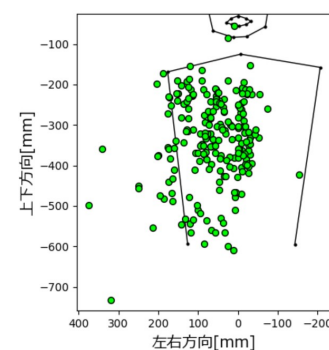
図 6 女性の顔の他の部位に関する分布



(a) 男性 fc

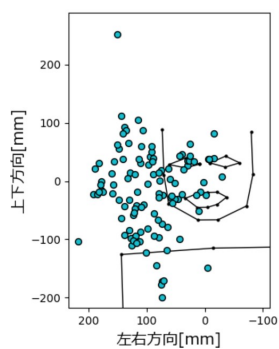


(b) 男性 tk(u)

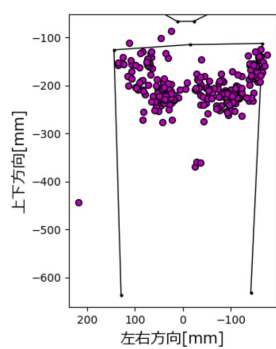


(c) 男性 tk

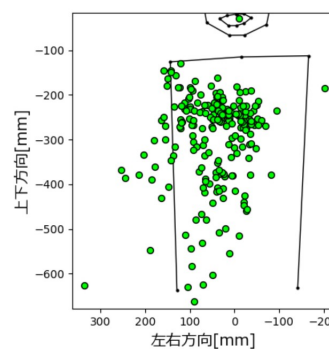
図 7 男性の顔の周りと胴体に関する分布



(a) 女性 fc



(b) 女性 tk(u)



(c) 女性 tk

図 8 女性の顔の周りと胴体に関する分布

異側の〈惜しい〉・〈ずるい〉, 図 3(b)と図 4(b)「lf(k)」と図 5(b)と図 6(b)「er」の〈ろうあ①〉, 図 3(c)と図 4(c)「lf(g)」の〈おいしい①〉, 図 5(a)と図 6(a)「mf」の〈謙虚〉が該当)は辞書データで接触部位が掌や手の甲と設定されているが実際の動作では指の部分の接触を行っていることから, 接触の表現に幅があると考えられる. そのため座標を抽出する代表部位の選び方についてさらなる検討が必要である.

四角形のマーカーが表す音節 (図 1(b)「uf(kmk)」男性の〈忘れる〉, 図 4(a)「lf(h)」女性の〈虎〉, 図 5(a)と図 6(a)「mf」の〈悪い①〉, 〈古い〉, 〈予定〉, 〈平気〉, 〈失敗①〉)が該当)は接近タイミングが始点と設定されているが, 実際の動作では始点は手が身体部位から離れた位置にあり, 動きの途中で付近を通過するように表現されている. これらの音節に関しては, 手と身体部位の距離が最短となる時点から座標を抽出し直すことによって身体部位周辺に分布するようになると考えられる.

ひし形のマーカーが表す音節 (図 1(c)「uf(m)」男性の〈盲〉・〈まちがう②〉, 図 5(b)と図 6(b)「er」の〈うわさ〉)が該当)は接近タイミングが動きの全体と設定されているが, 実際の動作では身体部位から離れた空間も利用した大きな動きで表現されている. これらの音節については, 動きの軌跡の重心を用いているが, 手と身体部位の距離が最短となる時点を用いるなど別の手法の検討が必要である.

×印のマーカーが表す音節 (図 2(b)「uf(kmk)」女性の〈涼しい〉, 図 3(b)と図 4(b)「lf(k)」の〈宣伝〉, 図 5(c)と図 6(c)「nk」の〈豊か〉, 〈脳卒中〉)は接近タイミングが始点や終点であるが, 実際の動作では全体的に身体部位から離れた位置で表現されている. これらの音節に関しては, 身体動作が容易な方向へ位置が変化した異音であると考えられる.

星形のマーカーが表す音節 (図 3(b)「lf(k)」男性の〈歯をみがく〉)は取扱 CL であり, 手の位置よりも取り扱う仮想の物体が存在する位置が重要であると考えられる. このような取扱 CL の扱い方に関しては今後の検討課題としたい.

図 7(a), (c)と図 8(a), (c)の「fc」, 「tk」は身体の中央と同側を用いて表現され, 図 7(b)と図 8(b)の「tk(u)」は身体の中央, 同側, 異側を用いて表現されるように, 広範囲に渡る部位についても分布が密集する箇所の特徴が見られる. また「tk(u)」や「tk」に関して, 〈注意〉と〈軍〉のように胸の中央と同側で対立すると考えられる音節が存在する. そのためこれらの位置の音素の考え方については今後の検討課題としたい.

音素の分布が他の音素の分布と重なっている部分が存在するが, 手話では位置の音素のみが異なるミニマルペアとなる音節の組み合わせが少ないため, 分布の重なりが許容されると考えられる.

4. おわりに

本研究では言語学に基づいて設定された手の位置の音素とろう者によって実際に表現された手話動作が持つ物理量との関係性を明らかにするために, モーションキャプチャで計測することで得られた高精度な 3 次元動作データと音素が記述されている辞書データを基に分析を行った. 結果として, 言語学的に設定された位置の音素が 3 次元空間でも対応する身体部位の周辺で表現されていることが明らかとなった.

今後は考察で扱った音節の事例も踏まえて, 座標を抽出する代表の部位や接近タイミングの設定の改善や音素や音節の定義に関する詳細な分析を検討したい. また手型や動きの音素に関する分析も検討したい.

謝辞 本研究は科研費 JP17H06114, JP20K11861 の助成による.

参考文献

- [1] Hara, D. A Complexity-Based Approach to the Syllable Formation in Sign Language. Chicago, IL: The University of Chicago, 2003, Ph.D. dissertation.
- [2] 日本手話研究所 (編). 日本語一手話辞典. 全日本ろうあ連盟, 2006, 2206p.
- [3] Watanabe, K.; Nagashima, Y.; Hara, D.; Horiuchi, Y.; Sako, S.; Ichikawa, A. "Construction of a Japanese Sign Language Database with Various Data Types". Communications in Computer and Information Science book series, 2019, Vol.1032, p. 317-322.
- [4] 西牧樹生, 堀内靖雄, 原大介, 黒岩真吾. "日本手話における手の位置の音素に関するモーションキャプチャによる分析". 人工知能学会研究会資料 言語・音声理解と対話処理研究会, 2021, vol. 91, p. 31-36.