

手話習得のための動作認識に基づく事前学習システムの開発

渡辺 雄大¹ 佐野 渉二² 中沢 実²

概要：我々は多くのことを学習したい欲求はあるが、学習することを意図すると学習が継続しない人も多い。これは、学習する対象が難しいと感じていることが理由の1つと考えられる。本研究では、意図した学習は継続しがたいことを考慮した事前学習を行う方法を考える。本稿では、手話を対象として、学習者が手話を学習することを意図しない中である程度の手話を学習できる事前学習システムを提案する。提案システムでは、手話とは関連がない動作を行う中でも手話の動きを認識した場合、その動きに対応する手話を音声で発する。これは、気付かれなかったり、理解できないことでも何回も繰り返す中で、動きと音声結びつくことになるという仮説を立て、その仮説に基づいて実装している。手話の学習とは伝えずに、けん玉とヨーヨーを行わせた評価実験において、提案システムによる事前学習の効果で手話テストでの正答率が向上し、意図しない中で手話の学習を一定程度行えることを示した。

Development of Pre-Learning System using Motion Recognition for Sign Language

YUDAI WATANABE¹ SHOJI SANO² MINORU NAKAZAWA²

1. はじめに

我々には、多くの知識や技能を学びたい欲求がある。ビジネスで利用するため、多くの人とコミュニケーションをとるため、人生を豊かに過ごすためなど理由は様々であるが、語学、金融、経済、プログラミング、音楽、スポーツなど多くのことを学びたい人は多い。しかし、必要に迫られていない中で、学ぶ意欲が高くない場合は継続しない人も多い。何かを学習する上ではある程度の困難は伴うものであり、学習することを考えることでその大変さを感じ、学習意欲をなくすことが継続しない理由の1つとして考えられる。仮に、習得したい知識や技能を登録するスマホアプリがあるとすると、多くの項目を登録する人がいても、実際に多くのことを学習する人は少ないことが推測される。

一方で、学校の授業は決められたカリキュラムに沿って行われ、学生は授業内容を決められないため学習意欲が低い人もいるが、教員の授業運営により、長い時間をかけて

ゆっくり学ぶこともあり、多くのことを学習できる。しかし、学習意欲が高くないと、学習の理解度や学習量は多くないことが多い。また、多くの学習支援システム [1], [2] が開発されているが、学習意欲が高いことを前提にしている。

このように、何かを学習するときには学習意欲が大きく関係していると考えられる。このため、我々は学習することを意識することでその困難さを感じるのであれば、学習すること自体を意図することなく学習を行えば、多くのことを学習できると考える。先に述べた習得したい知識や技能を登録するスマホアプリにおいて、登録した知識や技能が日常生活を送る中でいつの間にか習得できるのであれば、多くの人が利用するものと期待できる。そこで、本研究では学習することを意図しない中で学習する方法について考える。ただし、例えば、英語学習であれば、流暢に話せるようになればよいが、学習すると意図しない中で実用的なレベルまでに習得させるのは難しい。このため、本研究では、学習すると意図しない中で、一定程度の学習を行うことで、その後の学習に対する事前知識をつけたり、これによりモチベーションを高めたりするなど、学習のきっかけを作る事前学習を行うことを目的とする。

学習対象には先に挙げたように多くのものがあるが、本

¹ 金沢工業大学大学院工学研究科
Graduate School of Engineering, Kanazawa Institute of Technology

² 金沢工業大学工学部
College of Engineering, Kanazawa Institute of Technology

稿では手話の学習について考える。手話は、難聴者やろう者（以下、ろう者と記す）とコミュニケーションをとるために必要であり、ろう者と会話するために手話がよく用いられている。このため、手話を学習する人、学習したいと考える人が多いと考えられるため、手話の学習を扱うことにした。また、手話学習のために多くの手話学習支援システム [3], [4] も開発されている。しかし、これらは手話を学習する意欲がある人のためのシステムであり、多くの人は手話で話せたらよいとは感じて、必要に迫られない中では手話を学習をする人は少なく、手話学習支援システムを利用するに至らないことが推測される。身近にろう者がいたり、ろう者と会話を行う機会があったりするとき手話を学習するようになると考えられるが、必要になってから学ぶのでは習得するまでに時間がかかる。このため、日常生活の中で気付かないうちに手話を少しでも学ぶことができれば、手話を特別なものではなく、より身近なものであることと感じられることが推測される。これにより、手話を学習するための敷居が下がり、さらには手話学習の意欲にもつながることを期待され、さらには、上記の手話学習支援システムを利用することで、手話で会話ができるようになることにつながることも考えられる。

そこで、本稿では手話を対象とした事前学習システムを提案する。提案システムでは、前もって設定した手話の動きを認識した場合にその手話の意味を音声で発する。これは、音声に気付くか分からなくても、気付いてもその意味が分からないことでも、何回も繰り返すことで動きと音声結びつくことになるという我々が立てた仮説に基づいている。この仮説が正しければ、ある動作をしながら提案システムによる音声を感じることににより、意図しない中で手話の事前学習を行うことができるものと考えられる。けん玉、ヨーヨーを行う中で、意図せず手話を習得させるための評価実験を行った結果、提案システムにより動作を認識することで音声が発せられる回数が多い手話の方が、手話テストの正答率が高く、提案システムにより手話の学習が一定程度行われたと考えられる。

2. 関連研究

事前学習は、学習効果を高めることができ、多くの研究がされている。例えば、大学の授業用としても、学生の視聴反応を取得し、それを対面授業に活かすためのシステム [2] や、PBL (Project Based Learning) についてマンガで説明する事前学習教材 [5] が開発されている。また、学習内容をデジタル教材を用いて授業前に学習し、授業において演習や発展的な内容を学ぶ反転授業もよく行われている。例として、小学生を対象に反転授業を分析した結果、成績が下位群の児童にも一定の知識が定着されたことも確認されている [6]。これらにより、事前学習の効果が高いことが示されているが、これらは当然ながら学習対象を把握

した上で学習しており、学習する意図がない学習については対象外である。

また、学習にはモチベーションを維持することが重要であり、学習者のモチベーションを考慮した研究も多く行われている。学校の授業に限定しなければ、例えば、ピアノ学習においてモチベーションを考慮したピアノ学習支援システム [7] が開発されている。ミス許容度を導入し、ピアノのどの鍵を打鍵しても正しい音高の音が出力される 1 番許容度が高いモードから、徐々に許容度を下げた 7 つのモードを設け、段階的にミス許容度を下げながら学習する。これにより、成功体験を得ながら学習できるように学習者のモチベーションを維持しやすくしている。学習対象が明確である学習においてはミス許容度を考慮することは有効であると考えられるが、本研究は、学習することを意図しない学習を想定しているため、学習中のモチベーションは関与しないため、利用することはできない。

さらに、人が意識しない中でその人に影響を与える研究も行われており、一例として、プライミング効果を利用し、関心事へ無意識に気付きを与えるシステム [8] が開発されている。HMD (Head Mounted Display) で投影する映像の片隅に提示した情報により、実世界上の関連情報に対する行動に影響を与えることを評価実験により示している。無意識に気付きを与える点では本研究と似ているが、学習対象が動作を伴うものの学習については映像だけを用いることは難しい。ただし、本稿で対象とする手話以外で、利用可能な学習対象もあると考える。

一方で、手話を学習するための支援システムも多く開発されている。例えば、Leap Motion [3], [4] や Kinect [9], [10] により学習者の手の動きを認識しながら、手話学習を支援するシステムも開発されており、自主学習を促すための機能 [3] や、ゲーム [9] により学習効果を高めているシステムもある。しかし、これらの手話学習支援システムは手話を学ぶモチベーションが高くないと手話学習支援システムを利用するに至らないことが考えられる。本稿では手話の学習を意図しない中で手話の学習を行うが、学習内容は限定的となる。このため、本稿で提案する事前学習システムにより手話のある程度学び、それにより手話に興味を持たせたり、モチベーションを高められる場合は、従来の手話学習支援システムを有効に利用されることにつながると考えられる。

3. 手話のための事前学習システム

3.1 想定する事前学習システム

何かを初めて学習するときはとても難しい内容のように感じたことも、少し学習すると実際はそれまで難しいことではなかったことを体験する人も多いことが推測される。例えば、プログラミング言語を初めて学ぶときにはとても難しいことを行うように感じて、少し学ぶことで少なくとも



図 1 対象とする手話

も学習以前よりは難しいと感じることが少なくなった人も多いことが推測される。これは、学習内容の難易度が把握できることが大きいと考える。このため、事前学習することで、知識が増え、その後の学習をしやすくなることに加え、学習対象がそれほど難しいものではないことに気付くこともあると考えられる。さらには、少し学習することで学習意欲の向上につながることも期待できる。

ここで、プログラミング言語や英語などの学習は、スマホアプリケーションを作りたい、機械学習を行いたい、多くの人と話したいなどのモチベーションがあって学習を始めることが多いが、日常生活において、なじみや接点が少なく、学習の機会がないが学ぶことが大切なものもある。その1つとして、手話が考えられる。手話は、学びたいと考えている人が多いことが推測されるが、ろう者が周囲にいない中で学習する人は多くないかもしれない。しかし、必要になったときに学習し始めるのでは、ろう者とスムーズに会話できるようになるまで学習に時間がかかる。このため、ある程度事前学習することで、それが手話学習のきっかけになることが期待できる。ただし、学習するモチベーションがない中で学習させることになるため、学習の敷居が低い方法が必要である。本稿では、学習することを意図しない中で一定程度の学習を行う方法を考える。なお、意図しない中での学習であるため、実際に利用する際には、学習期間は数週間から数か月程度と長期間となりうることをは想定している。

3.2 研究対象

本稿では、手話の事前学習を対象とする。学習することを意図させずに学習を行わせるには、学習対象を気付かせないためにも学習者には行動を限定せずに、自由に行動させた中で行えることが望ましい。しかし、行動を限定せずに行わせる場合、学習者の動作を認識することが難しくなる。このため、本稿では事前学習の効果を評価することを目的とし、学習者の動作を認識しやすくするため、移動することなく、手話と似通った動作があるけん玉とヨーヨー

を行っているときに限定する。また、対象とする手話についての詳細は3.3節で記すが、けん玉やヨーヨーを行う中で動きが似通った動作の手話、および、違う動作の手話の合わせて11種類の手話を対象とする。

3.3 提案システム

従来の学習方法は、学習者が学習することを意図するものであり、自主的に初めて学習するものの場合、学習が継続しないことが多い。一方で、本研究では、学習者が学習することを意図しない中で学習を行う方法を考えるが、学習する意図がない中で実用レベルまでの学習を行うことは困難である。そこで、本研究では学習をしやすくするために、学習する意図がない中で学習内容をある程度学習できることを目標とする。本稿では、手話を対象とし、手話を意図しない中で学習できる事前学習システムを提案する。提案システムでは、カメラを用いて学習者の動作を認識し、手話のジェスチャーと同じ動作を行った際に手話の意味をスピーカーから発する。例えば、けん玉やヨーヨーをする中で、集めるという手話の動作を行った際にスピーカーから“集める”と発する。学習者は、音声で発せられていることに気付かなかったり、気付いていても何のことかわからないことが想定されるが、これを繰り返すことで、手話とはわからなくても、動作と手話の意味が結びつくようになるという仮説を立て、その仮説に基づいて、提案システムを考案した。仮説が正しければ、後に音声を手話に関することであると分かったときに、いくつかの動作とその手話の意味が結びついていれば、その分だけ手話を学習できていることに相当すると考える。

本稿で対象とする手話を図1に示す。“風邪”、“幸せ”、“飲む”、“集める”、“こんにちは”、“熱”、“おはよう”、“羨ましい”、“うるさい”、“分からない”、“さようなら”の11種類である。これらは、手話で会話するとき日常で使うものであったり、けん玉やヨーヨーの動作に近かったりする動作である。けん玉やヨーヨーの動きに近い動作は、“集める”、“おはよう”、“羨ましい”である。例えば、“集

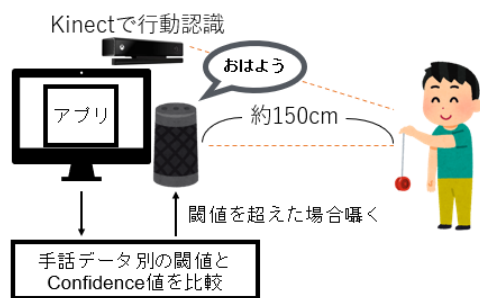


図 2 システム構成

める”の手話は、けん玉で糸の揺れを抑えるときの動作や、ヨーヨーを巻く動作と近く、“おはよう”の手話は、ヨーヨーを肩の位置から落とす動作に近い。

3.4 提案システムの設計と実装

提案システムの設計と実装を行った。提案システムのシステム構成を図 2 に示す。動作を撮影するカメラには、Kinect for Windows V2 を使い、学習者から約 150cm 程度離れた位置で学習者の動作を撮影して 25 点の骨格情報を取得し、デスクトップ PC 上で学習者の動作を認識する。認識した動作があらかじめ学習させてある手話の動作と一致すると判断するとき、スピーカから動作に合わせた音を発する。スピーカもデスクトップ PC の近くに設置する。

開発環境は Windows 上の Visual Studio であり、Windows Presentation Foundation を使用し、プログラミング言語は C# である。Kinect Studio を用いて Kinect で撮影を行うことにより、深度、色、骨格情報、音などを取得し、Visual Gesture Builder において Random Forest アルゴリズムを用いて図 1 に示す 11 種類の手話の動作を機械学習する。

学習者に手話の動作をさせ、学習データを生成すると動作の認識率は高まるが、本研究では学習者の学習している意図がない中で学習させる必要があるため、学習者に手話の動作をさせると気付かれやすい。このため、本稿では第 1 著者が手話の動作を行い、学習データを生成させることとした。機械学習を行う際の設定パラメータを表 1 に示す。表 1 の教師データフレーム数は第 1 著者が手話の動作を行って学習する際のものである。“風邪”、“飲む”、“おはよう”、“羨ましい”、“分からない”の手話は、口元に手を持ってくるといった似通った動作であるため、学習させるフレーム数を多くした。

次に、Kinect を用いて取得した動作データと学習データとの類似度を求め、その類似度があらかじめ設定した閾値を超えると手話の意味を音声で発する。表 1 には設定した手話の閾値も示す。第 1 著者が手話の動作を行ったときに認識率が 80% 程度、誤認識が 20% 以下になるように閾値を調整し、第 1 著者がテストしてみたところ、平均で 82% の認識率であった。また、手話の動作を短時間に複数回認識し、連続して音声を発すると気付かれやすいことを考慮

表 1 設定パラメータ

手話の動作	教師データのフレーム数	動作類似度の閾値
風邪	2256	0.20
幸せ	1431	0.23
飲む	1287	0.50
集める	1283	0.40
こんにちは	1161	0.70
熱	1310	0.50
おはよう	4154	0.40
羨ましい	2662	0.55
うるさい	1307	0.15
わからない	2410	0.40
さようなら	1145	0.20

し、手話の動作を認識した後 30 秒間は手話の動作を認識させないようにした。動作認識時に発する音声は、機械音で 11 種類の手話に対して作成した。

4. 評価実験

4.1 評価実験環境の構築

評価実験において、被験者に手話の学習を行っていると感じさせにくくするために、けん玉、ヨーヨー対する一人称研究 [11], [12], つまり、一人称視点でけん玉、ヨーヨーの技能を向上させるため、それぞれの習得方法に基づいて技能を向上させる研究を行うと装うことにした。このため、図 3 に示すように、けん玉、ヨーヨーの練習を振り返りやすくするためのシステムを構築した。このシステムでは、以下の機能を有する。

- メモ
 - ー 練習振り返りに、練習を通じて感じたこと、発見したこと、学んだことなどを書き込む。
- タイマー
 - ー 練習開始時間から経過した時間を表示する。
- 目標達成時間の設定
 - ー 計画立てて練習を行えるようにすることと、練習時間の目安にするために、目標時間を設定する。
- 通知
 - ー 設定した時間になったら通知する。
- BGM
 - ー 被験者には集中するために流すものと伝えるが、提案システムが発する音声を目立たせなくするためもある。

評価実験においては、図 2 に示すように、Kinect、スピーカから約 150cm 程度離れた場所で、被験者にけん玉、ヨーヨーを行わせた。提案システムにおいて動作認識時に発する音声は、BGM より少し大きい音量にすることで、BGM に紛れて聞こえる程度とし、感覚としてどこからともなく囁かれていると感じられるものにした。

4.2 評価実験の方法

被験者 A、被験者 B、被験者 C、被験者 D、被験者 E とした 20 歳から 23 歳の被験者 5 名に対して、以下の手順



図 3 集中力管理アプリケーション

で、すべて同じ部屋で評価実験を行った。

- (1) 一人称研究として、けん玉、ヨーヨーをそれぞれ 1 日 10 分間ずつ行わせ、けん玉、ヨーヨーの技能を向上させる。
- (2) 手話の事前学習を行っている事情を説明し、手話のテストを行う。
- (3) 手話の動作を行わせ、提案システムの動作認識精度を確認する。

手順 (1) においては、けん玉とヨーヨーをそれぞれ 1 日 10 分ずつ、計 20 分の練習を行った。けん玉検定の 10 級から 7 級まで使用される技、ヨーヨーのベーシックトリックの Set 1 で使用される技を行わせた。評価実験に慣れさせるためや、提案システムにより発する音声に気付かれにくくするため、1 日目には、提案システムにより発する音声の回数が少なくなるように、手話の動作と比較的似通っていないものが多いけん玉、2 日目に、手話の動作と似通っているものが多いヨーヨーとし、評価実験は連日で実施した。評価実験を行っている様子を図 4 に示す。図 4 (左) はけん玉、図 4 (右) はヨーヨーをそれぞれ行っている。評価実験では、けん玉、ヨーヨーを行う中で感じたこと、発見したこと、学んだこと、コツなどをメモ機能を用いて入力させる。さらに、一人称研究の評価実験であることを印象付けるためにも、けん玉を 2 日間練習した後にけん玉検定、ヨーヨーを 2 日間練習した後にヨーヨーのベーシックトリック検定を行うことを事前に伝え、それぞれ実施した。

手順 (2) では、手話を学習するためのシステムであることを伝えた後に、手話の学習度合いを確認し、提案システムの有用性を評価するためにも、手話のテストを行った。手話のテストでは、けん玉、ヨーヨーの動作から動作認識回数が多くなると予測される“おはよう”、“集める”、“さようなら”の手話、動作認識回数が少なくなると予測される“幸せ”、“風邪”、“羨ましい”の手話、本稿では対象としてない“こんばんは”の手話の 8 種類に対して、手話のイラストを見せ、その意味を答えさせた。

手順 (3) については、評価実験ではけん玉、ヨーヨーの動



図 4 評価実験の様子

表 2 評価実験中のメモ入力回数についての結果

被験者	1 日目 (けん玉)	2 日目 (ヨーヨー)
被験者 A	4	3
被験者 B	2	2
被験者 C	2	2
被験者 D	2	2
被験者 E	2	2

表 3 評価実験後に行った技能テストの結果

被験者	けん玉検定	ヨーヨーのベーシックトリック検定
被験者 A	10 級失敗	3 つ中、3 つ成功
被験者 B	10 級失敗	3 つ中、2 つ成功
被験者 C	10 級失敗	3 つ中、3 つ成功
被験者 D	10 級失敗	3 つ中、3 つ成功
被験者 E	10 級失敗	3 つ中、2 つ成功

きをさせる中で手話の動作を行っていることを認識するため、一見、手話の動作をしたように見える場合でも手話と提案システムにより認識しないことも想定される。また、第 1 著者が行った手話の動作に対して学習データを作成したため、被験者が手話の動作を行う場合の認識率は低くなることが想定されるため、評価実験として妥当な精度で手話の動作を認識できているかを測定する。具体的には、本稿で対象とする 11 種類の手話について、“風邪”、“幸せ”、“飲む”、“集める”、“こんにちは”、“熱”、“おはよう”、“羨ましい”、“うるさい”、“わからない”、“さようなら”の順で 5 回ずつ手話の動作を行わせた。

以降で記す手順 (1)、手順 (2)、手順 (3) については、上記のものを指す。

4.3 評価実験の結果

4.3.1 手話学習についての気づかれ度合い

本研究では学習していると意図していない状況での学習の効果を測定することが目的であり、被験者には手話の学習を行っていることを言わずに、けん玉、ヨーヨーに関する一人称研究の評価実験として行わせた。このため、評価実験を通じて手話の学習と気づいていたかを確認する。

表 2 にそれぞれの被験者が評価実験中にメモとして記録した状況についての結果を示す。被験者 A は、1 日目 4 回、2 日目 3 回の計 7 回で、それ以外の被験者は 1 日目、2 日目で 2 回ずつの計 4 回のメモを入力していた。また、

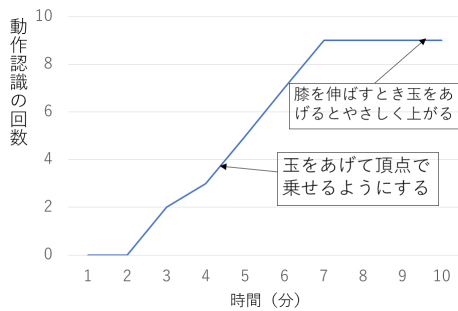


図 5 被験者 B についての動作認識回数の推移とメモ内容

被験者 B がけん玉を行っているときの提案システムによる動作認識の回数の推移とメモの内容を図 5 に示す。横軸は評価実験開始からの時間、縦軸は提案システムによる動作認識の回数を示す。被験者 B は 4 分過ぎと 9 分後半でメモしており、他の被験者も同じような状況であった。メモされた内容は、ほとんどの被験者で膝を上手に使うことや手首のスナップが大切であるなどのコツであった。一方で、被験者 C のみ、“あふれる”や、“さようなら”と言われたという提案システムにより発せられた音声についてもメモを行っていた、しかし、手順 (2) で評価実験について説明した時の反応から手話の評価実験であることは気づいていなさそうであった。

次に、けん玉検定、ヨーヨーのベーシックトリック検定の結果を 3 に示す。短時間ということもあり、すべての被験者においてけん玉検定は 10 級の合格には至らなかったが、ヨーヨーのベーシック検定では、すべての被験者が 2 つ以上のトリックを成功しており、技能の練習をしっかりと行っていたことが分かる。

以上より、評価実験において被験者は、手話学習のための評価実験を行っていることに気付いていなかったと判断できる。

4.3.2 提案システムの性能に関する評価実験の結果

手順 (2) でけん玉、ヨーヨーの練習を行っているときの提案システムによる動作認識に関する評価実験の結果を表 4、手順 (3) でけん玉、ヨーヨーを行った後に、被験者に手話の動作を行わせた際の動作認識に関する評価実験の結果を表 5 に示す。表 4、表 5 では、提案システムで推定する動作が実際にはどのような動作を行っていたかの回数について、すべての被験者のものを合算した回数を示しており、緑色に塗った部分が提案システムによる動作認識が正しく行われたことを表す。なお、提案システムが認識しなかった動作については記載していないため、表 5 でもそれぞれの動作が手順 (3) で示した試行回数に達していない。実際の動きについては、評価実験中に撮影したビデオカメラ映像から第 1 著者が判断した。

学習データは、第 1 著者が行った手話の動作で作成しており、被験者と身長、体格、腕の長さや骨格情報が異なるため、認識精度はある程度低下することは想定していたが、

概ね精度よく動作認識されていた。手順 (3) で被験者に手話の動作を行わせた場合の動作認識精度は F 値で 0.79 であり、けん玉、ヨーヨー練習時の動作認識精度は F 値で 0.73 であり、当然ながら前者の方が動作認識精度は高い。

実際の動作と違っていても、提案システムが認識した動作により対応する手話の音声を発するため、動作認識の精度は提案システムにおいては重要であるが、気付かれにくさの観点からは、多少の間違ひはカムフラージュとなるため、評価実験における動作認識の結果は許容範囲内と判断する。

表 6 には、手順 (2) でけん玉、ヨーヨーの練習を行っているときに、提案システムにより動作認識が行われた時に音声が発せられた回数を示す。提案システムで誤認識した場合でも、その誤認識した動作に対応する音声が発せられるが、誤認識の場合でも実際に発せられた音声をカウントした。表 6 中の「け」はけん玉、「ヨ」はヨーヨーを表す。すべての被験者に対する動作認識回数を合算した回数では、“風邪”が 2 回、“幸せ”が 1 回、“こんにちは”、“うるさい”が 0 回で動作認識回数が少なく、“集める”が 50 回や、“おはよう”と“さようなら”がそれぞれ 16 回と動作認識回数が多かった。被験者や認識された動作に差はあるが、手順 (1) で示した想定通り、1 日目より 2 日目の方が提案システムによる動作認識が行われていることが分かる。けん玉、ヨーヨーの練習時間がそれぞれ 10 分間であることを考えると、けん玉を行っているときは 1、2 分に 1 回、ヨーヨーを行っているときは 40 秒から 1 分に 1 回手話に対応付けられた音声が発せられたことになる。実際に提案システムを利用する場合は、1 日に動作認識による音声が発せられても 1、2 回程度で、数週間から数か月かけて学習することを想定するものの、評価実験においては、音声が発せられる間隔が比較的短時間なため、被験者によっては違和感を感じたかもしれないが、音声が発せられた回数としては概ね想定通りであった。

4.3.3 提案システムの有用性に関する評価実験の結果

けん玉、ヨーヨーの練習を終え、事情を説明した後に行った手話テストの結果を表 7 に示す。手順 (1) で示した想定通り、手話のテストの問題内で“おはよう”、“集める”、“さようなら”は動作認識回数が多く、“幸せ”、“風邪”、“羨ましい”の動作認識回数は少なかった。“考える”、“こんにちは”は、本稿では手話の対象としてない言葉である。表 7 において、被験者が手話の意味を正しく答えた場合は○、間違えた場合は×とし、被験者が答えた解答を示す。

被験者に見せた手話のイラストから正解が分かったものもありうるが、提案手法による動作認識により音声が発せられた回数と、手話テストの正解率には相関があることが分かる。つまり、表 6 の提案システムの動作認識により音声が発せられた回数について、4 回以上の言葉は正答率が 80%、1 回～3 回の言葉は 63%、本稿で対象としていない

表 4 提案システムによる動作認識と実際の動作に関する結果 (けん玉, ヨーヨー実行時)

		提案システムで推定する動作											再現率
		風邪	幸せ	飲む	集める	こんにちは	熱	おはよう	羨ましい	うるさい	わからない	さようなら	
実際の動作	風邪	6	0	0	0	0	0	0	0	0	0	0	1.00
	幸せ	1	2	2	0	6	0	0	4	0	2	5	0.09
	飲む	0	0	7	0	0	0	0	0	0	0	0	1.00
	集める	0	1	0	18	0	0	0	0	0	0	3	0.82
	こんにちは	0	0	0	0	10	1	0	0	0	0	0	0.91
	熱	3	0	0	0	0	3	0	0	0	0	1	0.43
	おはよう	4	0	1	0	0	0	18	0	0	0	2	0.72
	羨ましい	0	0	0	1	0	0	0	16	0	0	0	0.94
	うるさい	0	0	0	0	0	0	0	0	10	0	0	1.00
	わからない	0	0	0	0	0	0	0	0	0	9	0	1.00
	さようなら	0	0	0	0	0	0	0	0	0	0	12	1.00
適合率		0.43	0.67	0.70	0.95	0.63	0.75	1.00	0.80	1.00	0.82	0.52	0.73

表 5 提案システムによる動作認識と実際の動作に関する結果 (手順 (3))

		提案システムで推定する動作											再現率
		風邪	幸せ	飲む	集める	こんにちは	熱	おはよう	羨ましい	うるさい	わからない	さようなら	
実際の動作	風邪	7	0	0	0	0	0	0	0	0	1	0	0.88
	幸せ	1	4	2	0	9	0	0	2	0	1	7	0.15
	飲む	0	0	10	0	0	0	0	0	0	0	0	1.00
	集める	0	0	0	23	0	0	0	0	0	0	0	1.00
	こんにちは	0	0	0	0	12	0	0	0	0	0	0	1.00
	熱	1	0	0	0	0	3	0	0	0	0	0	0.75
	おはよう	3	0	1	0	0	1	21	0	3	0	1	0.70
	羨ましい	0	0	0	0	0	0	0	19	0	0	0	1.00
	うるさい	0	0	0	0	0	0	0	0	14	0	0	1.00
	わからない	0	0	0	0	0	0	0	0	0	12	0	1.00
	さようなら	0	0	0	0	0	0	0	0	0	0	15	1.00
適合率		0.58	1.00	0.77	1.00	0.57	0.75	1.00	0.91	0.82	0.86	0.65	0.79

表 6 提案システムによる動作認識により発せられた音声の回数

	A		B		C		D		E	
	け	ヨ	け	ヨ	け	ヨ	け	ヨ	け	ヨ
風邪	0	0	2	0	0	0	0	0	0	0
幸せ	0	0	0	0	0	0	0	0	1	0
飲む	0	1	0	0	0	0	0	0	0	0
集める	1	4	1	8	4	8	0	12	3	9
こんにちは	0	0	0	0	0	0	0	0	0	0
熱	1	0	4	0	0	0	0	0	3	0
おはよう	1	3	0	4	0	3	3	1	1	0
羨ましい	0	0	0	2	0	0	0	0	0	0
うるさい	0	0	0	0	0	0	0	0	0	0
わからない	1	0	0	0	0	1	0	0	0	0
さようなら	2	5	2	0	0	5	0	1	0	1

言葉の正答率は 55% である。このため、提案システムにより手話の学習が一定程度行われたと考えられる。

4.4 アンケート結果

評価実験において意図しない中で手話学習の機会が得られたが、これが手話を学習するきっかけとなりうるかを判断するため、1 が“そうは思わない”、2 が“あまり思わない”、3 が“変わらない”、4 が“少し思う”、5 が“そう思う”の 5 段階で回答させた。アンケート結果を表 8 に示す。

アンケート結果では、手話への興味はあるが、手話について新たに調べることはないことが示されている。つまり、手話への興味に対する否定的な回答が被験者 A のみであり、提案システムにより手話学習のきっかけとなりう

ることが示されたが、これにより手話を学習しようという意欲が沸くことは少なそうである。評価実験の時間が 1 日あたり 10 分を 2 日間と短かったことが考えられるが、被験者に意識させない中で手話の学習行わせた評価実験であるので、意識的に回答するアンケートにはその感情的な変化が表面化されなかったことも考えられる。

5. 考察

5.1 今後の展望

評価実験の結果、手話を学習する意図がない中で手話の学習を一定程度行われたと考えられるが、評価実験が 1 日あたり 10 分間の 2 日間という短時間であったこと、けん玉、ヨーヨーを行わせた後で行わせた手話のテストの問題数が少なかったことから、評価実験結果の信頼性が十分あるとは言えない部分がある。このため、今後は評価実験の期間を長くして、評価実験を再度行う予定である。また、学習の達成度を確認する手話テストに対しても学習度合いがより分かるテストを行うことを考えている。

さらに、提案システムにより得られる学習効果が手話テストにより表面化しない部分もあると考えられる。つまり、本稿での評価実験後のアンケート結果では示されなかったが、提案システムの有無により興味や学習意欲に差が生じ、その後、手話支援システムを用いて手話を学習するときの

表 7 手話テストの結果

被験者	おはよう	考える	集める	幸せ	さようなら	こんばんは	風邪	羨ましい
被験者 A	× (楽しい)	○	○	× (おはよう)	○	× (はじめまして)	○	○
被験者 B	○	○	○	× (おはよう)	○	× (はじめまして)	○	○
被験者 C	○	○	× (こんにちは)	× (考える)	○	○	× (難しい)	○
被験者 D	○	○	○	○	○	× (はじめまして)	○	○
被験者 E	× (楽しい)	○	○	× (こんにちは)	× (おはよう)	× (さようなら)	× (難しい)	× (風邪)

表 8 アンケート結果

被験者	手話に興味を持てたか	手話について調べようと思ったか
被験者 A	1	3
被験者 B	3	4
被験者 C	4	2
被験者 D	4	1
被験者 E	3	3

継続性や学習効果に影響を及ぼすことも考えられる。これは、提案システムはしっかりと学習する前の事前学習という位置付けであり、本格的に手話の学習を行わないと事前学習の効果を示せないことがある。このため、評価実験を再度行う際には、その後、手話の学習を行わせて、提案システムの有用性を評価する予定である。

5.2 実用化に向けた考察

本稿では、けん玉、ヨーヨーを行いながら、意図しない手話の学習を行わせることを行った。現状では、けん玉やヨーヨーを日常的に行っている人に対し、手話の事前学習を行わせるために利用できる。つまり、けん玉やヨーヨーの練習会において、それらを行っているときの姿勢を解析するために、カメラを設置することは可能と考える。ただし、暗にでも手話を学びたい人はけん玉やヨーヨーを行う人以外の方が多いと考えると、応用範囲を広げる必要がある。本稿では、実装の都合上、1 台の Kinect を利用したが、一か所にとどまらず、移動する学習者を対象とする必要があるため、実用化するためには、複数台カメラにより学習者を認識 [13] しながらの動作認識を行う必要があると考えられる。移動する学習者に対しては、装着型小型カメラ [14] や加速度センサ [15] による動作認識を行うことも考えられる。

6. おわりに

本稿では、学習する意図がない中で学習を行う事前学習システムを提案した。Kinect により学習者の動作データを取得し、骨格情報から手話の動作を認識したら、手話の意味を音声で発する。けん玉、ヨーヨーを行う中で手話を学習させるシステムを実装し、評価実験を行った結果、提案システムが動作を認識し、その意味が発せられた手話に対して理解度が高いことから、提案システムによりある程度の手話を学習できたものと考えられる。

今後の課題として、本稿よりも長い時間の評価実験を行うこと、提案システム使用後に手話を学ぶことにより学習効率が上がるかを検証することにより、提案システムが事

前学習システムとして機能するかを評価することが挙げられる。また、けん玉、ヨーヨー以外の動作、行動に対しても対応できるように、提案システムを拡張する予定である。

参考文献

- [1] 安藤大岳, 大場 みち子: 自主学習継続のための学習支援システムの構築, 情報処理北海道シンポジウム 2014, No.A-07, pp.20-23 (2014)
- [2] 芝崎順司: オープンコンテンツを利用した事前学習のための視聴反応モバイルシステムの開発と利用評価, 教育メディア研究, Vol.2, No.2, pp.25-37 (2015)
- [3] 森本正志, 川除慎吾, 加藤雅弥, 田端孝成, 加藤秀康, 永井敦, 竹内健人: モーションセンサを用いた指文字学習支援アプリケーション, 情報処理学会論文誌: デジタルコンテンツ, Vol.8, No.2, pp.28-40 (2020)
- [4] Schioppo, J., Meyer, Z., Fabiano, D., Canavan, S.: Learning Sign Language in a Virtual Environment, Proceedings of the CHI Conference, No.1422, pp.1-6 (2019)
- [5] 高橋聡, 高橋 B. 徹, 吉川厚: PBL の事前学習教材としてのマンガケース教材およびその設問セットの提案, 科学教育研究, Vol.41, No.2, pp.116-130 (2017)
- [6] 稲垣忠, 佐藤靖泰: 家庭における視聴ログとノート作成に着目した反転授業の分析, 日本教育工学会論文誌, Vol.39, No.2, pp.97-105 (2015)
- [7] 竹川佳成, 福家悠人, 柳英克: モチベーションを考慮したピアノ学習支援システム情報処理学会論文誌, Vol.57, No.4, pp.1193-1206 (2016)
- [8] 磯山直也, 寺田努, 塚本昌彦: ユーザの関心事へと引き込みを行なう常時映像閲覧システム, ヒューマンインタフェース学会論文誌, Vol.17, No.1, pp.39-52 (2015)
- [9] Kamnardsiri, T., Hongsit, I., Khuwuthyakorn, P., Wongta, N.: The Effectiveness of the Game-Based Learning System for the Improvement of American Sign Language using Kinect, The Electronic Journal of e-Learning, Vol.15, Issue. 4, pp.283-296 (2017)
- [10] Mustafa, E., Dimopoulos, K.: Sign Language Recognition using Kinect, Proceedings of the 11th International Conference on Artificial Intelligence and Soft Computing (ICAISC 2012), pp.394-402 (2012)
- [11] 諏訪正樹, 堀 浩一: 特集「一人称研究の勧め」にあたって, 人工知能学会誌, Vol.28, No.5, pp.688 (2013)
- [12] 出口正倫, 佐野渉二: 振り返り機能を有する筋力トレーニング支援システムの構築, 平成 29 年度北陸地区学生による研究発表会, No.E-4-2, p.66 (2018)
- [13] 森大樹, 内海章, 大谷淳, 谷内田正彦, 中津良平: 非同期多視点画像による人物追跡システムの構築, 電子情報通信学会論文誌 D, Vol. J84-D-2, No.1, pp.102-110 (2001)
- [14] Wu, E., Yeo, H., Quigley, A., Koike, H., Kitani, K.: Back-Hand-Pose: 3D Hand Pose Estimation for a Wrist-worn Camera via Dorsum Deformation Network, Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST2020), pp.1147-1160 (2020)
- [15] 渡邊光, 村尾和哉, 望月祐洋, 西尾信彦: 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO 2016) 論文集, Vol. 2016, pp. 1677-1684 (2016)