

VR 空間におけるスペイン語対話学習システムの提案

齊藤真生^{†1} コルドバユキコ^{†1} 櫻井淳^{†1}

概要: 我が国では、外国人労働者の増加などに伴い、英語をはじめとした中国語、韓国語やスペイン語などの外国語習得の必要性が高まっている。その学習においては、実際の教師を相手とした会話が理想的であるが、時間的かつ経済的な制約を受けるため、スマートフォン端末にて手軽に音声対話が可能なアプリが注目されている。しかし、その対話相手は機械であるため、学習者が会話の臨場感を得られないことや、質問に対して長時間考えることを許容する原因となり、人間同士の円滑なコミュニケーション技能を身に付けることは困難である。そこで、本研究では、ヘッドマウントディスプレイを用いた VR 空間におけるスペイン語の対話学習システムを提案し、その有効性を検証する。

1. はじめに

我が国では、外国人労働者の増加などに伴い、英語をはじめとした外国語習得の必要性が高まっている。そうした状況に対し、文部科学省では、2020 年度より小学校からの英語科目の教育を義務付ける[1]など、英語教育の一層の充実を図っている。一方、中国語、韓国語やスペイン語などのその他の外国語は、主として大学と一部の高等学校で教育されている現状である。文部科学省の調査[2]によると、外国語科目を開設している高等学校は、2001 年と 2018 年を比較すると、中国語が 424 校から 497 校、韓国・朝鮮語が 163 校から 342 校、スペイン語が 84 校から 96 校といずれも増加傾向にある。今後の国際化に伴って、多言語の学習者はますます増えていくと考えられる。

外国語の学習においては、実際の教師を相手とした会話が理想的であるが、時間的かつ経済的な制約を受けることから、安価かつ手軽に外国語を学習する手段としてコンピュータ支援型語学教育 (CALL: Computer Assisted Language Learning) システムが活用されている。これが登場し始めた 1990 年代の頃は、HTML ファイルを用いた Web 教材[3-5]や CD-ROM 教材[6]を活用し、ドリル型の練習問題やリスニングによる学習が主体であった。2000 年以降になると、語学の 4 技能のうち、話す・聞くの技能を重視した教育が重要視され、それらを支援するための音声情報処理技術を用いた CALL[7-9]が広く研究されるようになった。これらでは、学習者の音声と教師音声の発音や韻律を比較して評価する手法が提案されている。さらに、最近では、クラウドサービスや人工知能の技術発展により、音声認識や対話制御などの対話に必要な機能が向上し、ロボットを活用した対話システムの研究[10]や、スマートフォン端末を用いて手軽に音声対話で外国語を学習するアプリ[11,12]が登場している。しかし、これらの対話相手は機械であるため、学習者が会話の臨場感を得られないことや、質問に対して長時間考えることを許容する原因となる。こうした状況に

おいて、学習者が人間同士の円滑なコミュニケーションをとるための技能を身に付けることは困難である。この解決策として、CG キャラクタを対話相手として設定し、その表情を時間的に変化させることによって学習者にプレッシャーを与える方法が提案[13]され、一定の有効性が示されている。ただし、この研究では、ロボットで表現された簡易の CG キャラクタが使用されるため、人に近いデザインや表現に変えることで得られる効果は明らかにされていない。

そこで、本研究では、実際の人物との円滑なコミュニケーション能力を身に付けるための外国語の学習環境を構築することを目的とし、ヘッドマウントディスプレイ (以下、HMD) を用いた VR 空間におけるスペイン語の対話学習システムを提案する。なお、対象言語はスペイン語とするが、他の言語にも応用可能なシステムを構成する。まず、第 2 章では、本研究で提案するシステムの開発環境、画面構成、使用機器、処理フローと会話パターンについて述べる。次に、第 3 章では、実証実験として本システムの音声の認識精度、処理時間と活用効果を評価する。そして、第 4 章において、システムの改善に向けた試行内容を記述する。最後に、第 5 章にて本研究を総括する。

2. システムの開発

2.1 開発環境

本研究では、スペイン語の学習者に対して、自学自習による対話の練習を支援するために、VR 空間におけるスペイン語の対話学習システムを開発する。本システムの開発環境を表 1 に示す。ゲームエンジンには、3 次元モデルを実写の人物に似せるため、グラフィック性能が高い Unreal Engine を採用する。このエンジンは、Blueprint とよばれる専用のプログラミング言語を使用し、CG キャラクタの動作や画面遷移などの制御が行える。また、CG キャラクタを構築するため、まず、オープンソースの Make Human を用いて人体の基礎となるモデルを生成する。次に、そのモデルに対して色付けを行うために、Substance Painter を使用し、

顔や服などのリアルな質感のテクスチャを作成する。そして、CG ソフトの Blender を用いて目のまばたきや口の動きなどのアニメーションを追加する。一方、音声認識系の処理は、各種クラウドサービスのライブラリが充実しているプログラム言語の Python を使用する。また、ユーザの発話データとゲームエンジンとを連携するため、XAMPP を用いて JSON データを HTTP 通信する方式を取り入れる。

2.2 画面構成

本システムのシステム画面を図 1 に示す。本システムは、VR 上の CG キャラクタが質問する内容に対して、学習者がマイクを通じて返答内容を発話することで会話シナリオが進む構成とする。CG キャラクタが質問を行った後、タイムバーを表示させることにより、学習者にプレッシャを与える仕組みを取り入れる。これは、実際の対話場面において違和感のない時間を考慮し、質問から返答までにかかる時間を 5 秒間に設定した。また、CG キャラクタは、会話文に合わせた口の動きのアニメーションをそれぞれ作成することで、違和感のない動きとなるように工夫した。

2.3 使用機器

本システムの動作環境として、図 2 に示すように、HMD の Oculus Rift S とその付属コントローラ、マイク、HMD の動作要件を満たした GeForce GTX 1070 の GPU を搭載したゲーミング PC を使用する。これにより、VR 空間上で CG キャラクタと対話が可能となる。また、HMD の付属コントローラは、対話を開始する最初のタイミングのみで使用し、その後は自動で対話が進む構成とした。マイクは、HMD に付属されているが、周りが雑音の環境であっても会話音声を拾うことができるように、単一指向性のサンワサプライ製 USB ヘッドセット（型番：MM-HSU07BK）を用いた。

2.4 処理フロー

本システムの処理フローを図 3 に示す。まず、CG キャラクタの質問に対する学習者の発話データを入力データとして、Google Speech Recognition を使用してその音声をテキストデータに変換する。次に、そのデータを Watson assistant に送信し、予め用意した学習モデルをもとに返答内容の会話文を取得する。このとき、学習者の発話データが適切な場合は正解と判断し、その内容に見合った会話文を生成する。一方、不適切な場合は不正解と判断し、理解できなかったことを表す会話文とする。そして、その会話文を Watson Text to Speech で音声に変換する。最後に、Unreal Engine 上の CG キャラクタに対して、会話文の音声と、それに対応したアニメーションを動作させる。これを繰り返すことで会話シナリオの一連の流れを実現する。

2.5 会話パターン

本研究では、大学の語学科目としてスペイン語を履修している学生が学習する範囲を想定し、財団法人日本スペイン協会が主催するスペイン語技能検定 6 級の内容を基準に会話パターンを作成する。これは、挨拶、自己紹介や現在

表 1 開発環境

名称	用途
Unreal Engine	3次元向けのゲームエンジン
Make Human	人体製作用の3次元CGソフト
Substance Painter	3次元モデルのテクスチャ作成ソフト
Blender	3次元モデルの編集ソフト
XAMPP	HTTP通信用のソフト
Python	音声認識系処理のプログラミング言語



図 1 画面構成



図 2 使用機器

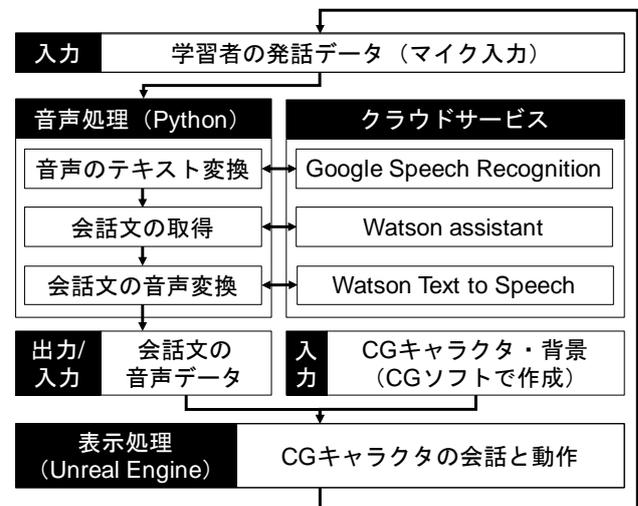


図 3 処理フロー

表 2 会話パターン

No.	CG キャラクタの質問	学習者の返答 (正解例)	CG キャラクタの返答
1	¿Qué estás haciendo ahora? あなたは何をしていますか?	<u>Estoy</u> estudiando español. 私はスペイン語の勉強をしています。	Estoy hablando Español. 私はスペイン語を話しています。
2	¿De dónde eres? どこ出身ですか?	<u>Soy</u> de Tokyo. 私は東京出身です。	Soy de Barcelona. 私はバルセロナ出身です。
3	¿Dónde está tu libro de español? あなたのスペイン語の教科書はどこにありますか?	Mi libro de español <u>está</u> en mi casa. 私の教科書は家にあります。	No lo tienes ahora. 私はそれを持っていません。
4	¿Dónde vives? どこに住んでいますか?	<u>Vivo en</u> yokohama. 私は横浜に住んでいます。	Vivo en Tokyo. 私は東京に住んでいます。
5	¿Hay alguien en casa? 家の中に誰かいますか?	<u>Está</u> mi madre. / <u>no hay</u> nadie. お母さんがいます。 / 誰もいません。	¡Mándale mis saludos! お母さんによろしく伝えてください。
6	¿Qué quieres comer? あなたは今なにを食べたいですか?	<u>Quiero comer</u> carne. 私はお肉が食べたいです。	¡Yo también! 私も食べたいです。
7	¿Cuántos años tienes? 何歳ですか?	<u>Tengo</u> veinte años. 私は 20 歳です。	Tengo veintiuno años. 私は 21 歳です。
8	¿Te interesa la computadora? あなたはこのコンピュータに興味はありますか?	<u>Si, me interesa.</u> はい、そのコンピュータに興味があります。	A mi me aburre. 私は退屈です。
9	¿Dónde está tu madre? あなたのお母さんはどこにいますか?	Mi madre <u>está</u> en casa. 私のお母さんは家にいます。	Vale, muy bien. いいですね。
10	¿A dónde vas a ir hoy? 今日はこれからどこへ行くつもりですか?	<u>Voy a</u> la Universidad. 私は文教大学に行くつもりです。	Yo voy de compras. 私は買い物に行きます。
11	¿Cómo te llamas? 名前は何ですか?	<u>Me llamo</u> saitou. / <u>soy</u> saitou. / <u>Mi nombre</u> es saitou. 私の名前は齊藤です。	Encantada de conocerla, me llamo yuri. はじめまして、私はユリです。
12	¿Qué día es hoy? 今日は何曜日ですか?	<u>Hoy es</u> miércoles. 今日は水曜日です。	Tengo clase de Español. 今日はスペイン語の授業があります。
13	¿Quién es tu profesor de español? あなたのスペイン語の先生は誰ですか?	Mi profesor <u>es</u> Terada. / es Terada. / <u>se llama</u> Terad. 寺田という名前です。 / 寺田さんです。	Mi profesor es sakurai. 私の先生は、櫻井です。
14	¿A dónde vas en las vacaciones? どこで休暇を過ごしますか?	<u>Voy a</u> mi pueblo. 自分の実家に行きます。	Yo voy a Hokkaido. 私は北海道に行きます。
15	¿Te gustan las verduras? あなたは野菜が好きですか?	<u>No me gustan</u> las verduras. / <u>Me gustan</u> las verduras. 私は野菜が好きではありません。 / 私は野菜が好きです。	A mi también. / A mi tampoco. 私も好き。 / 私も嫌い。
16	¿Dónde trabaja tu padre? あなたのお父さんはどこで働いていますか?	Mi padre <u>trabaja en</u> una peluquería. 私のお父さんは床屋さんで働いています。	Mi padre trabaja de Ingeniero. 私のお父さんはエンジニアです。
17	¿A qué hora sales de casa? あなたは何時に家を出ますか?	<u>Salgo a</u> la una y media. 私は午前一時半に家を出ます。	Está bien, nos vamos さあ、行きましょうか。
18	¿Qué estudias? あなたは何の勉強をしていますか?	<u>Estudio</u> español. 私はスペイン語の勉強をしています。	Yo estudio italiano 私はイタリア語の勉強をしています。
19	¿Qué tiempo hace hoy? 今日はどんな天気ですか?	<u>Está</u> soleado. 今日は晴れています。	¡Gracias! ありがとう。
20	¿Puedes ir conmigo a ver una película? 私と一緒に映画を見に行ってくれませんか?	<u>Puedo ir</u> contigo. 私はあなたと映画を見に行けます。	¡Qué bien! / ¡Qué lástima! それはよかった。 / それは残念。

形の文章など、基本的な会話ができる程度の難易度であり、1 級から 6 級までである中の最も入門レベルの内容である。会話内容は、市販テキストの対象の級に該当する文法や単語を使用して作成を行った。また、文教大学のスペイン語の授業の担当講師にヒアリングし、1 セメスタ分で学習する範囲の文法と単語を使用することとした。その理由として、この授業は 6 級レベルの能力習得に向けて構成されているが、一部の単語は扱わないことを配慮したためである。作成した会話パターンの一覧を表 2 に示す。会話パターンは、CG キャラクタの質問内容、学習者の返答内容とそれ

に対する CG キャラクタの返答内容を 1 つのセットとする。学習者の返答内容については、下線部分の単語がすべて含まれている場合に、正解と判定する。

本システムでは、全 20 種類の会話パターンに対して、No.1~5, 6~10, 11~15, 16~20 の 5 パターンごとに分割することで 4 種類のシナリオを作成する。そして、これらのシナリオの中から 1 種類を選択し、5 個の会話パターンを体験する流れとした。なお、学習者の発話内容や、その正誤の結果は、CSV ファイルでテキストデータとして出力し、体験後に確認して結果を振り返るために使用する。

3. 実証実験

3.1 実験概要

本章では、本システムの有効性を検証するための実証実験を実施する。実験内容は、図4に示すように、実験Ⅰ．音声の認識精度の評価、実験Ⅱ．システムの処理時間の評価、実験Ⅲ．システムの活用効果の評価の3種類とする。実験Ⅰと実験Ⅱでは、学習者が違和感なく本システムを使用できるかの観点で評価し、実験Ⅲでは、実際に本システムを活用した際の学習効果を評価する。

3.2 実験Ⅰ．音声の認識精度の評価

(1) 実験方法

本実験では、学習者の発話データを入力データとし、Google Speech Recognition による音声の認識精度を評価する。被験者は、日常からスペイン語を使用しているネイティブスピーカーの被験者 A (女性) と、大学で2年間スペイン語を勉強して日常会話が可能な被験者 B (男性) の2名とする。実験方法として、各被験者が表2に示す20種類の会話パターンを発話し、その内容のとおりテキストに変換されているかを確認する。

(2) 実験結果

実験結果を表3に示す。被験者 A は20回中19回、被験者 B は20回中18回で認識に成功し、平均の認識率は92.5%であった。認識に失敗した原因として、動詞の es と固有名詞の名前が結合されたことや、名詞が異なる単語として認識されたことがある。そのため、一つずつの単語の区切りを明確にし、流暢に話すことができれば、ほぼ100%の精度で認識可能といえることがわかった。

3.3 実験Ⅱ．システムの処理時間の評価

(1) 実験方法

本実験では、CG キャラクタの質問に対して5秒間で学習者が返答内容を発話した後、図4の緑枠部分に示す音声処理を行い、CG キャラクタが会話文を返答するまでに要した時間を評価する。被験者は、実験Ⅰと同様に、被験者 A と被験者 B の2名とする。実験方法として、各被験者が表2に示す20種類の会話パターンを発話し、それぞれに対して2名の返答時間の平均値を算出する。また、CG キャラクタの質問、学習者の返答、それに対するCG キャラクタの返答の単語数とその合計数をそれぞれ記述し、単語数に対して返答時間が変化するかを確認する。

(2) 実験結果

実験結果を表4に示す。20個の会話パターンの平均値として、合計単語数が11.40個に対して処理時間は3.39秒であった。また、個別の結果に着目すると、最小処理時間はNo.14の単語数14個に対して2.78秒、最大処理時間はNo.15の単語数11個に対して4.01秒であった。これらのことから、処理時間は単語数によって影響するのではなく、クラウドサービスとの通信時間が影響している可能性が高い。

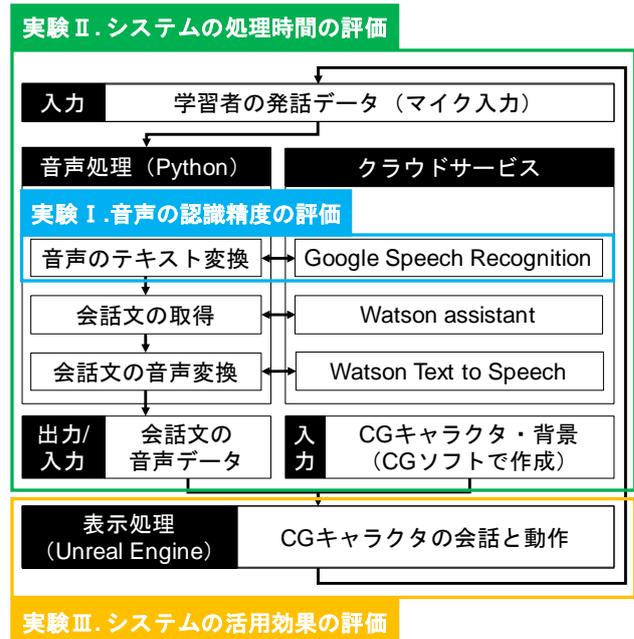


図4 実験計画

表3 音声の認識精度の評価結果

対象	会話数	音声認識の結果		
		認識成功	認識失敗	認識率
被験者 A	20	19	1	95%
被験者 B	20	18	2	90%
平均	20	18.5	1.5	92.5%

表4 システムの処理時間の評価結果

No.	CG 質問 単語数	返答 単語数	CG 返答 単語数	合計 単語数	処理 時間
1	4	3	3	10	3.41 秒
2	3	3	3	9	3.75 秒
3	6	8	4	18	3.09 秒
4	2	3	3	8	3.24 秒
5	4	3	3	10	3.31 秒
6	3	3	2	8	3.08 秒
7	3	3	3	9	3.41 秒
8	4	3	4	11	3.49 秒
9	4	5	3	12	3.66 秒
10	6	4	4	14	3.77 秒
11	3	3	6	12	3.27 秒
12	4	3	4	11	3.61 秒
13	6	4	4	14	2.86 秒
14	6	4	4	14	2.78 秒
15	4	4	3	11	4.01 秒
16	4	6	5	15	3.42 秒
17	6	6	4	16	3.77 秒
18	2	2	3	7	3.47 秒
19	4	2	1	7	3.38 秒
20	7	3	2	12	3.12 秒
平均	4.25	3.75	3.40	11.40	3.39 秒

また、3～4 秒程度で返答が返ってくることから、実際の人間との会話のように、違和感なく対話できると考えられる。

3.4 実験Ⅲ. システムの活用効果の評価

(1) 実験方法

本実験では、本システムの活用による学習に与える効果を評価する。被験者は、大学の1セメスタ分のスペイン語授業の履修を終えて2セメスタ目を履修中の学生を対象に、被験者C（女性）と被験者D（男性）の2名とした。実験方法として、表2の会話パターンを用いて、被験者CはNo.1～10、被験者DはNo.11～20の10パターンずつをシステム上で体験し、その正否を判定する。実験回数は2回とし、1回目は事前対策を行っていない状況で実施する。また、2回目は、1回目で体験した内容の質問と返答の文章を紙で配布し、1日間の勉強期間を与えたうえで、1回目の翌日に同様の質問内容で実施する。

これに加えて、体験結果に対して学習効果に関するアンケートを行う。内容は、会話のスムーズさ、話しやすさ、緊張感、リアルさと学習効果の5種類とし、良い5・やや良い4・普通3・やや悪い2・悪い1の5段階で評価する。また、自由記述とヒアリングにより意見を聴取した。

(2) 実験結果

まず、正否の実験結果を図5に示す。この結果より、被験者Cは0から5、被験者Dは1から7に正解数が増加した。なお、被験者Cに関して、2回目は5回正解であるが、実際には8回で正解の文章を話していた。この3回の差分については、クラウド側の音声認識に失敗したことが原因であるため、発音やイントネーションが正しくなかったといえる。以上より、本システムを活用することで技能向上に一定の効果が得られたと考えられる。ただし、この実験では、正解の文章を提示して勉強期間を設けたことや、質問の出題順を固定にしたことから、本システムの学習による効果ではなく、自学自習の効果として結果に反映された可能性が高い。これに対しては、質問の出題順をランダムにすることや、本システムで繰り返し学習させることによって、さらにシステムの活用効果を検証できると考える。

次に、アンケート結果を図6に示す。会話のスムーズさ、リアルさや緊張感において高評価であるため、人間とのリアルなコミュニケーションが体験できていると考えられる。一方、話しやすさは、両者とも低い評価であった。この原因を被験者に確認したところ、実際の講師との対話の場合は、相手に合わせてゆっくりと話すなどの配慮があるが、本システムの会話スピードは一定であるため、ペースを合わせるために気持ちが焦ってしまうとの意見であった。これに対しては、質問に返答できない場合は会話スピードを落として再度質問を行うなどの工夫が考えられる。また、正しい単語を話しているにもかかわらず認識しない場合があったため、音声認識の正確さの点で評価が低くなった。これに対しては、現状はクラウドサービスを使用した音声

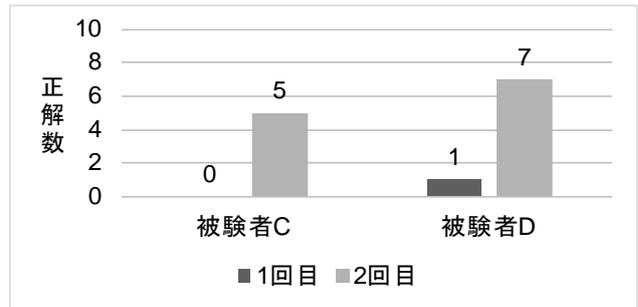


図5 システムの活用効果の評価結果

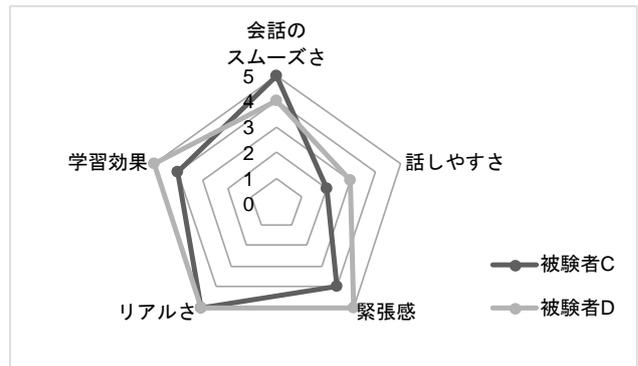


図6 システムの活用効果のアンケート結果

認識であるため、それを使用せずに音声認識処理を実装する必要がある。その関連研究として、日本人の音声データを収集し、深層学習などを用いて日本人の発話内容を認識する手法の可能性が提案[14,15]されているため、こうした手法をもとに開発することが考えられる。

以上より、本システムを用いることで、実際の人物とのコミュニケーションに近い練習が可能であり、本システムによる学習効果の一定の有効性を示せたといえる。また、よりリアルな会話に近づけるためのいくつかの改善点の案を得られたため、これらは今後の課題である。

4. システムの改善に向けた試行

本システムでは、リアルな対話を実現するために、CGキャラクターに対して、身振りや口の動きのアニメーションを設定している。特に、口の動きに関しては、各会話文に対して個々に設定しているため、会話パターンを増やすごとにアニメーションを作成する必要があり、作成コストを要する。そこで、実際の人物の動きをカメラで撮影し、その動きをトレースできれば作成作業を大幅に効率化できることから、既存手法を用いて身体と口のそれぞれの動きの自動化の試行を行った。各試行結果の詳細を記述する。

4.1 身体のリース

身体のリースには、3d-pose-baseline[16]の手法を使用する。これは、単眼カメラで撮影した人物に対して、OpenPose[17]で2次元骨格データを生成し、深度推定デー

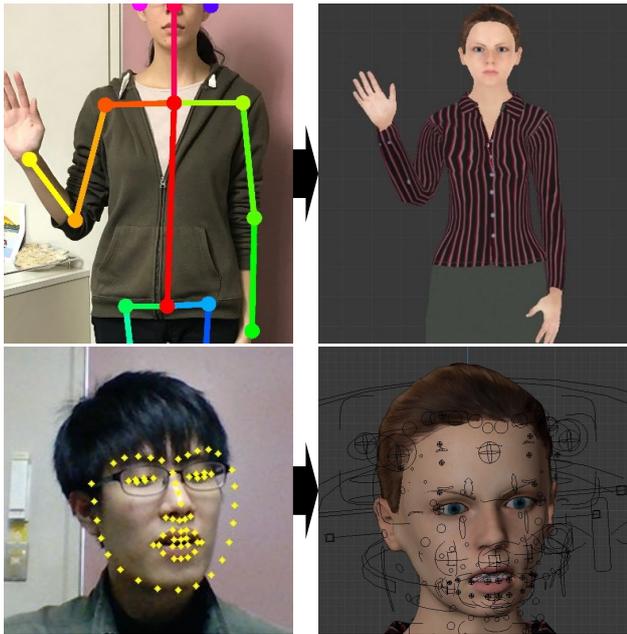


図 7 身体と顔の撮影とトレース結果

と組み合わせて 3 次元骨格データを生成する手法である。

試行結果を図 7 上部に示す。単眼カメラを使用しているため、奥行き軸の挙動が再現できない場合も見受けられたが、大まかな動きは再現することができた。そのため、この手法で動きを作成した後に CG ソフト上で微修正する方法であれば、十分に活用でき、コスト削減につながる。

4.2 顔のトレース

顔のトレースには、FacialMotionCapture[18]を使用する。これは、OpenCV にて単眼カメラ上の人物の顔の特徴点を抽出し、CG キャラクタ上に対応した特徴点を配置することで顔の動きを再現する手法である。

試行結果を図 7 下部に示す。カメラ上の顔の傾きや口の動きに合わせた CG キャラクタの動作を確認できたため、これを用いて人物が話す口の動きを再現可能と考えられる。ただし、CG キャラクタを自作したため、口の開き方に制限がないことや歯茎の位置が正常に表現されない問題があるため、実際の人物に近づけるための調整が必要である。

5. おわりに

本研究では、実際の人物との円滑なコミュニケーション能力を身に付けることを目的とし、VR 空間におけるスペイン語の対話学習システムを提案した。そして、3 種類の実証実験により本システムの有効性を確認した。実験 I ではほぼ 100%の精度でスペイン語の発話内容を認識できること、実験 II では実際の人間との会話のように違和感なく対話できること、実験 III では本システムを活用して一定の学習効果が得られることをそれぞれ示した。また、システムの改善に向けて CG キャラクタの動作の自動化を試行し

た。今後は、心拍センサやカメラなど IoT 機器と組み合わせるなどにより、学習者の反応によって CG キャラクタの動作を変えるなど、よりリアルな対話の実現を目指す。

謝辞 本研究の遂行にあたり、文教大学スペイン語講師の寺田裕子氏にご協力頂いた。ここに記し感謝の意を表す。

参考文献

- [1] 文部科学省. 新学習指導要領について. https://www.mext.go.jp/b_menu/shingi/chousa/shisetsu/044/shiryo/_icsFiles/afieldfile/2018/07/09/1405957_003.pdf, (参照 2019-12-23).
- [2] 文部科学省. 平成 29 年度高等学校等における国際交流等の状況について. https://www.mext.go.jp/a_menu/koutou/ryugaku/koukousei/_icsFiles/afieldfile/2019/09/19/1323946_001_1.pdf, (参照 2019-12-23).
- [3] 堀田英夫. CALL 教室でのスペイン語教育: HTML で作成した教材を利用して. コンピュータ&エデュケーション. 1999, vol.7, p.70-74.
- [4] Spelichal, M. and Inoue, S.. The Roundtrip Method via HTML File as a Language Learning Strategy. Japanese Journal of Educational Media Research, 2000, vol.6, no.2, p.41-49.
- [5] ビラールイリヤス. 語学用 Web 教材に汎用性をもたらす試み. コンピュータ&エデュケーション. 2001, vol.11, p.37-40.
- [6] 中條清美, 西垣知佳子, 内堀朝子. 英語初級者向け CALL システムの開発とその効果. 日本大学生産工学部研究報告 B 文系. 2005, vol.38, p.1-16.
- [7] 中川聖一. CALL と音声情報処理技術. 音声研究. 2005, vol.9, no.2, p.28-37.
- [8] 峯松信明, 富山義弘, 吉本啓, 清水克正, 中川聖一, 壇辻正剛, 牧野正三. 英語 CALL 構築を目的とした日本人及び米国人による読み上げ英語音声データベースの構築. 日本教育工学雑誌. 2003, vol.27, no.3, p.259-272.
- [9] 中川聖一, 牧野正三, 壇辻正剛. 音声言語処理技術を用いた語学学習システム. 日本音響学会誌. 2003, vol.59, no.6, p.337-344.
- [10] 秋本桃子, 阿部秀尚, 生田祐子, 森田武史, 山口高平. 教師業務ルール分析に基づく対話型ロボットを用いた発音練習の実装と評価. 情報教育シンポジウム論文集. 2018, vol.2018, no.26, p.185-188.
- [11] ジョイズ株式会社. TerraTalk. <https://www.terratalk.rocks/>, (参照 2019-12-23).
- [12] appArray 株式会社. SpeakBuddy. <https://www.speakbuddy.me/>, (参照 2019-12-23).
- [13] 鈴木直人, 廣井富, 千葉祐弥, 能勢隆, 伊藤彰則. 応答タイミングを考慮した英会話練習のための音声対話型英語学習システム. 情報処理学会論文誌. 2015, vol.56, no.11, p.2177-2189.
- [14] 大崎功一, 峯松信明, 広瀬啓吉. 日本人英語発声に観測される発音上の癖を考慮した音声認識. 電子情報通信学会技術研究報告. SP, 音声. 2003, vol.102, no.749, p.7-12.
- [15] 南條浩輝. 多言語音声の音声認識. 日本音響学会誌. 2018, vol.174, no.9, p.531-534.
- [16] Julieta, M., Rayat, H., Javier, R. and James J. L.. A simple yet effective baseline for 3d human pose estimation. The IEEE International Conference on Computer Vision (ICCV). 2017, p.2640-2649.
- [17] Zhe, C., Gines, H., Tomas, S., Shih-En, W. and Yaser, S.. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. CVPR. 2018.
- [18] "FacialMotionCapture". <https://github.com/jkirsons/FacialMotionCapture>, (参照 2019-12-23).