

# バイノーラル再生を用いた 英語学習における臨場感提示システムの構築

大城 澄香<sup>†1</sup> 井村 誠孝<sup>†1</sup>

**概要:** 近年, 日本において英会話能力は企業や教育の現場で重要視されている。また日本人が海外に行くことや, 外国人の来日も増え, 国際交流が盛んであることから, 日本では英会話に対する学習意欲が年々高まっている。このニーズに応えるために英会話教室や e-learning などがあるが, 英会話教室は時間の制約があり, e-learning では英会話教室ほどの臨場感が得られないといった問題が挙げられる。また, 臨場感を向上させるために VR 技術を用いた英語学習システムも存在するが, 聴覚刺激や, 主体的に学習に関わっていく点においてユーザに働きかける要素が欠けている。そこで本研究では, 臨場感を向上させるために, バイノーラル再生に着目する。これは, バイノーラル録音により得られる頭部の音響効果を再現した音源を聴取すると, その場に居合わせたかのような臨場感を再現できる技術である。視覚からの情報に連動した聴覚刺激が可能なシステムを構築し, 主体的な学習を実現する。

## Construction of Presence Providing System for English Learning Using Binaural Reproduction

SUMIKA OSHIRO<sup>†1</sup> MASATAKA IMURA<sup>†1</sup>

**Abstract:** In recent years, ability of conversation in English is important in companies and education sites. Also, since Japanese people go abroad and the number of foreign visitors come to Japan more frequently and international exchange is prosperous, the motivation for learning English conversation is increasing year by year in Japan. There are English conversation classes and e-learning to respond to this demand, but the English conversation class has time constraints, e-learning cannot afford a realistic feeling as much as the English conversation classroom. There is also an English learning system using VR technology to improve the presence, but there are lacking elements to encourage users in respect of auditory stimuli or subjects involved in learning subjectively. Therefore, in this research, we focus on binaural reproduction in order to improve the realism. This is a technology that can reproduce the realistic feeling as if it existed on the spot when listening to the sound source reproducing the acoustic effect of the head obtained by binaural recording. We construct a system capable of auditory stimulation linked with information from the vision and realize subjective learning.

### 1. はじめに

2016 年の TOEIC による調査によると, 現在の日本人の英語会話能力は, 72 ヶ国中 35 位と非常に低く [1], 過去 5 年で見ても低下傾向にある。日本人の英語力が低い理由として, 英語学習が成功した時の具体的なイメージが湧かないことが原因であると考えられる。したがって, 実際に英語を使う臨場感を提示することで学習への主体的な意欲を促進することが重要である。

学習方法を提案するにあたり, 学習理論に基づいて考える。学習理論とは, 学習の原因や仕組みを説明する心理学の理論で, 古典的学習理論と正統的周辺参加理論の二つがある。古典的学習とは, 個人が技術と知識を机に向かって主体的に習得していく学習プロセスのことであり, 正統的周辺参加とは実践コミュニティに参加することを通して学

ばれる知識と技能の学習プロセスのことを言う。また, 学習の中でも英語に着目した場合, ESL (English as a second language) と EFL (English as a foreign language) の二つの学習法がある。ESL とは学習者の環境に英語が使用されている状況を指し, EFL とは, 学習者の環境に英語以外の言語が使用されている場合の状況を指している [2]。つまり EFL 型教育を受けている我々にとっては, 環境が欠けている点が問題である。本研究ではこの環境の欠損を VR (virtual reality) 技術を用いて補うことで日本人の英語会話能力向上を図る。

また, 近年教育に VR を活用する試みが広くなされており, VR を用いた英語学習支援システム等 [3][4], 学習に臨場感をもたせることで没入感を与える学習方法もいくつか提案されている [5]。しかし, 従来のシステムは, 単語単位の学習であり, 実際に話すような会話形式ではない。また, 聴覚臨場感においても, 視覚に連動した会話等は再現されていない。

本研究では, 臨場感を向上させるために, バイノーラル

<sup>†1</sup> 関西学院大学理工学部  
School of Science and Technology, Kwansai Gakuin University

再生に着目する。また、会話形式を用いるために音声認識を行う。バイノーラル再生とは、得られる頭部の音響効果を再現した音源を聴取すると、その場に居合わせたかのような臨場感を再現できる技術である。視覚情報に連動した聴覚刺激が可能なシステムを構築し、主体的な学習を促進する。本稿では、臨場感を向上させるためのシステムの構築について述べる。

## 2. 提案手法

本研究では、バイノーラル再生を用いた英語学習における臨場感提示システムを提案する。体験者には、英語での対話環境において視覚的・聴覚的に没入できる空間を提示し、かつその会話に体験者が加わることができるシステムを構築する。会話の反応に応じて次のシーンに移る仕組みを作成し、ユーザが主体的に会話を行う体験を実現する。コンテンツの構成図を図1に示す。視覚的にも聴覚的にも臨場感のある空間を提示することによってシステムへの没入感が増すと考えられる。また、音声認識を用いた会話形式にすることによって、ユーザには主体性が必然的に求められる。

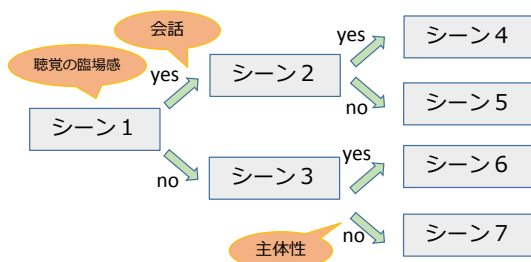


図1 コンテンツ構成  
Figure 1 Components of contents

システム構成を図2に示す。1)体験者はHMDとヘッドホンまたはイヤホンを着用する。2)全天球動画に写っている2名の会話の最後に体験者への問いかけを行う。3)問いかけに対して応答すると、その応答に応じた次のシーンに移り変わる。4)シーンが切り替わり、再び2)と3)を行う。これが繰り返されることによって会話が成り立つ。

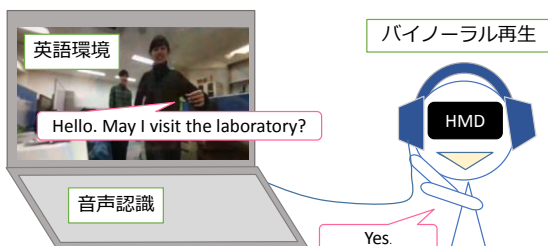


図2 システム構成  
Figure 2 Components of System

## 3. 臨場感に関する予備実験

### 3.1 実験内容・方法

臨場感におけるバイノーラル再生の有効性を明らかにするためにアンケートに基づいた主観評価実験を行った。実験では大学生9名(男:4人,女:5人)を対象に、バイノーラル録音とステレオ録音の音源を提示し、以下3つの項目において4件法で評価を行った。再生にはヘッドホンを用いた。1)どのくらい臨場感を感じたか。2)話に引き込まれたのはどちらか。3)どちらの音源の方が好きか。音源の再生順序はランダムにし、視覚からの情報の影響をなくするため実験は閉眼で行った。

### 3.2 使用音源

音源は、3種類の会話(2人)をそれぞれバイノーラル録音とステレオ録音した計6種類を使用した。バイノーラル録音の際には、頭部伝達関数を考慮するため本研究では自身の頭を使用した。録音した音源は音源ごとに違和感のないよう音量を一定にした。

### 3.3 実験結果・考察

1)の臨場感の評価において、バイノーラル再生とステレオ再生に差があるか否かを対応のあるt検定により評価した。検定の結果、 $p = 9.0 \times 10^{-11} < 0.01$ となったことから、1%水準で有意差が確認された。また、3種類の会話を1つずつ検定した場合も1%水準での有意差が見られた。この結果から、主観評価実験においてバイノーラル再生はステレオ再生と比べて臨場感が得られることが示された。

## 4. 実装システム

### 4.1 システム概要

2節で提案したシステムの開発を行った。プラットフォームは統合型のゲーム開発環境であるUnityを用い、シーンを3つ作成した。まず全天球カメラTHETA S(RICOH社)によって全天球動画を、シーンを包含する球体の内側にマッピングすることでシーンの背景とした。

次にバイノーラル再生を行うためにUnityのプラグインであるOculus Audio SDK1.0.1を使用した。これにより体験者は、話者の位置が考慮され、かつ視点の動きに合わせた音声を体感できる。話者の音声は入力した言語をmp3音源として出力するSound of textサービスを使用して生成した。実際に全天球動画を貼り付け、バイノーラル再生を行えるよう設定したものが図3である。



図3 視覚臨場感提示の例

Figure 3 Example of visual presence

音声認識は大語彙連続音声認識エンジン Julius を用いて行った。体験者の発話が「Yes」ならばシーン2へ、「No」ならばシーン3へ移り変わるようシステムを構築した。誤認識を防ぐため、認識単語辞書には「Yes」と「No」のみを登録した。

#### 4.2 コンテンツ内容

本システムに用いる会話内容について説明する。今回は理系学生に向けた内容を提案する。設定としては、外国人が研究室に入って来るところから会話がスタートし、研究内容について話をする。はじめに体験者が置かれる環境設定について用紙を用いて説明する(図4)。外国人はまれに状況と違った内容を話すため、体験者の聴く力が試される。今回は理系学生に向けた内容設定としているが、動画と音声があれば、様々なシチュエーションに応用することができる。会話の例を図5に示す。

- 現在の時刻 午前10時
- 今日の予定
  - 2限 (11:10~12:40) : 授業
  - 3限 (13:30~15:00) : 進捗報告
- 明日の予定
  - 昼休み (12:40~13:30) : 研究室ミーティング
  - 夜 (19:00~) : 飲み会
- 研究内容
  - バーチャルリアリティ

Figure 4 Experience setting

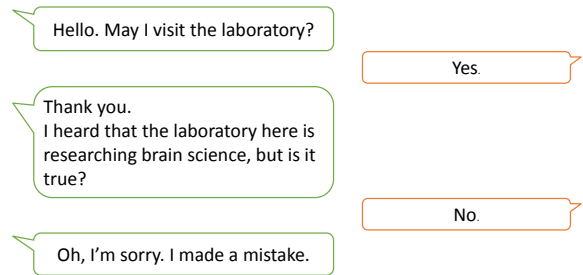


図5 会話例

Figure 5 Example of conversation

## 5. おわりに

本稿では、バイノーラル再生を利用した英語学習が主体的な学習に繋がると考え、バイノーラル再生が英語学習に及ぼす効果の検証について述べ、臨場感のある英語音声によるシステムの提案を行った。主観評価実験においてバイノーラル録音がステレオ録音よりも臨場感を感じられることを示し、聴覚刺激による臨場感があり、かつ会話形式による主体的な学習ができるシステムの作成を行った。

今後、体験者が高い臨場感を得られるよう HMD を用いて全天球動画を提示し、全天球動画に合わせたバイノーラル再生が行えるよう改善を行う。また、聴覚的な臨場感を向上させるため、音声として肉声を録音したものを使用し、最後にシステムの評価実験を行う予定である。

## 参考文献

- [1] <http://www.efjapan.co.jp/epi/> (2016.12.15)
- [2] 沖原勝昭: 英語教育における ESL と EFL の違いについて, 京都ノートルダム女子大学研究紀要, Vol. 41, pp. 69-80, 2011.
- [3] <https://devpost.com/software/house-of-languages> (2016.12.15)
- [4] 小島隆次, 竹鼻圭子, 楠見考: 視覚イメージを利用した英語前置詞 over の空間的意味学習における接触要素に関する実験的検討, 日本教育工学会論文誌, Vol. 31, Suppl., pp. 205-208, 2007.
- [5] 小島隆次, 米田英嗣, 竹鼻圭子, 森本雅博, 楠見考: 仮想空間を用いた英語前置詞教材の効果的利用, 日本教育工学会論文誌, Vol. 31, No. 2, pp. 219-228, 2007.