

# 発話内容を文字と関連画像で被面接者に重畳表示する AR 型面接の提案と有効性の検討

中村浩幹<sup>†1</sup> 中村亮太<sup>†2</sup> 上林憲行<sup>†1</sup>

東京工科大学<sup>†1</sup> 湘北短期大学<sup>†2</sup>

## 1. はじめに

入学試験や採用試験等における対面面接では、緊張や説明不足等の理由で伝えたいことが面接官に伝わらないことがある。そこで本研究では、面接官が AR ゴーグル (MOVERIO BT-200) を装着し、音声認識によって面接中に被面接者が発言する内容をすべて字幕で表示させた上で関連した画像も表示させることにより、対面状況で被面接者が面接官に伝えたいことを伝えられる手助けになるような面接支援を提案する。本稿では提案手法について説明し、提案手法に対して伝達感 (どの程度、話の内容が伝わってきたと感じたか) と伝達度 (実際にどの程度、話の内容がもれなく伝わっているか) [1] の 2 つの評価基準から有効性の検討を行う。

## 2. AR 面接

先行研究では、表情などが見えるといった理由でチャットなどの非対面よりも対面でのコミュニケーションの方が、伝達感が高いということが報告されている。しかし伝達度に関しては対面よりも非対面コミュニケーションの方が高い [1]。これはチャットのように文字で内容を伝えることで読み手が内容をより強く記憶した可能性がある。

著者らは先行研究から高い伝達感を得るためには対面という状況が必要であり、それに加えて高い伝達度を得るためには文字等を利用した視覚情報を聞き手に提示する必要があるのではないかと考えた。

## 3. 実験

### 3.1 AR 面接の構成

Google Cloud Speech API を用いて開発した音声認識アプリケーションによって面接中に被面接者が発言する内容をすべて字幕で表示させ、同時に関連画像も表示させる。



図1 アプリケーションの実行画面

図 1 の字幕と書かれている部分には被面接者の発話内容がすべて字幕化され、画像と書かれている部分にはキーワードに対応した画像が表示される仕組みである。例えば「私が学生時代に頑張ったことは、高校生の時に剣道の県大会で個人優勝したことです。」と被面接者が発話した場合、そのままの字幕が表示された上で個人優勝というキーワードに反応して、被面接者が当時個人優勝を達成した時の画像が表示される。今回はプロトタイプシステムであるため、キーワードと画像の登録方法については事前に被面接者からキーワードのテキストとそれに一致した画像を受け取り、音声認識アプリケーションに手動で登録している。面接官は AR ゴーグルによって図 1 の画面を被面接者に重ね合わせて面接を行う。なお被面接者も目の前に設置しているスマートフォンで図 1 の画面を確認することができる。また被面接者の発話内容の認識もスマートフォンのマイクで行っている。

### 3.2 実験方法

本実験は、通常面接と AR 面接の伝達度を比較することを目的としている。著者らは比較対象として計 2 回の面接実験を行った。1 回目は通

A proposal for an AR type interview that overlays the speech and related images to the applicant.

<sup>†1</sup> HIROKI NAKAMURA, NORIYUKI KAMIBAYASHI, Tokyo University of Technology

<sup>†2</sup> RYOTA NAKAMURA, Shohoku College

表1 伝達度の比較結果 (平均と標準偏差)

グループ	人数	平均値	標準偏差	平均値の標準誤差
通常面接	7	7.71	3.45	1.30
AR面接	7	11.29	2.63	0.99

常面接で、2回目はAR面接である。面接官は就職活動を終えた大学4年生と短大2年生のメンバーから通常面接で4名、AR面接で4名の計8名に行わせた。被面接者は短大1年生から通常面接で7名、AR面接で7名の計14名に行わせた。10分間で1対1の面接を実施し、面接官は「自己紹介」「成長を実感したエピソード」「最近の関心ごと」の3点について質問した。これらは実際の就職活動でよく聞かれる質問である。図2に実験環境を示す。



図2 AR面接実施環境

AR面接の実施環境としては、まず開発した音声認識アプリケーションを表示するスマートフォンが被面接者の目の前に設置される。被面接者は自分の発話内容が音声認識されると同時に、自分の発話内容の字幕やキーワードに応じた画像を確認することができる。そのアプリケーション画面をARゴーグルにもキャストして表示しているため、面接官も同時に確認できる仕組みとなっている。実験では各面接終了後に面接官に覚えている内容を用紙に書き出させた。

### 3.3 実験結果

面接終了後に面接官が書き出した内容を単語ごとに分解し、実際の面接内容と一致した単語をカウントすることで、通常面接とAR面接における平均単語数を比較した[表1参照]。通常面接は平均7.7単語であるのに対して、AR面接で

表2 伝達度の比較結果 (t検定)

	等分散性のためのLeveneの検定		2つの母平均の差の検定		
	F	有意確立	有意確立(両側)	平均値の差	差の標準誤差
等分散が仮定されている	0.609	0.45	0.050	-3.571	1.639
等分散が仮定されていない			0.052	-3.571	1.639

は平均11.3単語と有意に増加したことを確認した[表2参照]。先行研究から対面という状況であれば高い伝達感を得られることが判明していたが、AR面接はそれに加えて高い伝達度も得られることが分かった。

### 4. 考察

実験よりAR面接は伝達感の高い対面という状況でありながら、伝達度も通常面接より高いといえるだろう。面接官に実施したアンケート結果からもAR面接の方が話しやすく、理解もしやすくなるという結果が得られている。しかしながら被面接者にとっては話しにくいという意見があった。この主な原因としては、面接官がゴーグル上の文字や画像を見ている時に目が合わないことだった。ARゴーグルに表示するコンテンツの配置を工夫するなどして自然なアイコンタクトを可能にする必要がある。

### 5. おわりに

本研究は、対面面接において被面接者の発話内容に応じて文字と画像をリアルタイムにARゴーグルに表示することが可能なAR面接を提案した。実験結果よりAR面接は対面という状況から得られる高い伝達感と、通常面接より高い伝達度を得られることが示唆された。字幕だけでなく、キーワードに応じて被面接者の過去の画像が表示されるという仕組みに伝達度を高める大きな可能性があると考えられる。

### 参考文献

- [1] 杉本陽子(2010)「インターネットコミュニケーションと対面コミュニケーションにおける情報の伝わり方の差異についての意見書」[http://www.kantei.go.jp/jp/singi/it2/kaikaku/dai3/siryu3\\_2\\_2.pdf](http://www.kantei.go.jp/jp/singi/it2/kaikaku/dai3/siryu3_2_2.pdf) (参照2017-7-19)
- [2] Intelligent Technology's Technical Blog, 2017「Google Cloud Speech APIの精度を試す」<http://iti.hatenablog.jp/entry/2017/05/09/112051> (参照2017-11-7)