

# 運転支援を目的とした合成音声の印象比較

向 直人<sup>1,a)</sup> 山中 南実希<sup>1</sup> 高嶋 恵子<sup>1</sup> 藤掛 和広<sup>2</sup> 田中 貴紘<sup>2</sup> 金森 等<sup>2</sup>

**概要:** 近年, 65 歳以上の高齢運転者による死亡事故の割合は年々増加し, 平成 28 年には事故発生件数の 54.8 % を占めるなど, 高い水準で推移している. 高齢運転者の運転事故を防ぐには, 同乗者からの注意喚起や警告が有効だと考えられる. 本研究では, 合成音声を利用したエージェントによる運転支援に着目する. 合成音声で再現可能なピッチや抑揚などの感情表現は, ドライバの振舞いに対して影響を与える可能性がある. そこで, 特徴量の異なる合成音声を作成し, ドライバにもたらす運転支援効果の差を明らかにすることを本研究の目的とする. 合成音声の評価のため, ドライビング・シミュレータの走行映像を制作し, 一時停止や交差点などの状況に合わせて, 合成音声が生産される仕組みを構築した. 実験結果から, 感情を込めた合成音声は一部の被験者においては有効であるが, 強すぎる抑揚は却ってストレスをもたらす可能性があることがわかった.

**キーワード:** 合成音声, 感情表現, 運転支援, ドライビング・シミュレータ, エージェント

## Impression Comparison of Synthetic Voices for Driving Assistance

NAOTO MUKAI<sup>1,a)</sup> NAMIKI YAMANAKA<sup>1</sup> KEIKO TAKASHIMA<sup>1</sup> KAZUHIRO HUZIKAKE<sup>2</sup>  
TAKAHIRO TANAKA<sup>2</sup> HITOSHI KANAMORI<sup>2</sup>

**Abstract:** In these years, fatal accidents of people aged 65 or older are increasing yearly, and they account for 54.8% of the total number of accidents in 2016. A driving assistance by fellow passengers in a car is considered effective to reduce such accidents. In this work, we focus on driving assistance agents which speak synthetic voices. We assumed that emotional expression such as pitch and intonation of synthetic voices have an effect on driving performance. Thus, we aim at revealing a difference among some synthetic voices which have different features. In order to evaluate the synthetic voices, we created a video of a driving simulation which includes typical driving situations, e.g., stop sign. Experimental results indicated that a synthetic voices with emotion has limited effectiveness, and an excessive intonation has the potential to add extra stress to drivers.

**Keywords:** Synthetic Voices, Emotional Expression, Driving Assistance, Driving Simulator, and Agent

### 1. はじめに

総務省統計局の調査<sup>\*1</sup>によると, 近年, 交通事故の発生件数, 及び, 交通事故による負傷者数は減少傾向にある. しかし, 65 歳以上の高齢運転者による死亡事故は年々増加し, 事故発生件数の半数を占めていることが報告されている. 一般に, 高齢運転者は, 高齢であるほど, 操作不適や

前方不注意を要因とする死亡事故を起こしやすくなる傾向がある. 高齢運転者の運転事故を防ぐには, 同乗者が注意喚起や警告を行い, 運転者をサポートすることが有効と考えられる. しかし, 実際には高齢者が単独で運転する機会も多く, 常に同乗者のサポートが得られるわけではない. そこで, 本研究では, 合成音声を利用したエージェントによる運転支援に着目する. ここで, エージェントとは自律的なソフトウェアやロボットを指す.

近年, 合成音声技術の発展は目覚ましく, エージェントやロボットとの対話に関する研究が注目されている [1], [5], [6].

<sup>1</sup> 椋山女学園大学 文化情報学部 文化情報学科

<sup>2</sup> 名古屋大学 未来社会創造機構

<sup>a)</sup> nmukai@susigyma-u.ac.jp

<sup>\*1</sup> <http://www.stat.go.jp/>

文献 [4] では、合成音声の最新動向がまとめられており、これまで、音声の合成に広く用いられていた隠れマルコフモデル (Hidden Markov Model: HMM) に替わり、ディープニューラルネットワーク (Deep Neural Network: DNN) を用いた合成音声の研究が活発になっていることが報告されている。また、音声合成の実用化も進んでおり、株式会社エーアイは、コーパスベースの技術を採用して、喜び、怒り、悲しみなどの感情表現を伴う合成音声を生成可能な **AITalk** というソフトウェアを提供している [3]。AITalk の活用事例は幅広く、防災・広域放送システム、スマホ用アプリ、ロボットなどの分野で利用されている。本研究では、AITalk を利用し、運転支援に必要な注意喚起や警告などの合成音声を生成する。このとき、声の性別や抑揚などが異なる特徴量の合成音声を生成し、ドライバーにもたらず運転支援効果の違いを明らかにすることを目的とする。運転状況に合わせた合成音声を生成することで、「強い抑揚が注意効果を高める」、また、「女性の声で安心感を与える」、などの効果が期待できると考えた。合成音声の評価のため、ドライビング・シミュレータの走行映像を制作し、一時停止や交差点などのシチュエーションに合わせた合成音声で再生される仕組みを構築した。本稿では、予備実験として椋山女学園大学の学生を被験者として実施したアンケートの結果を分析し報告する。

本稿の構成を下記に示す。第 2 章では、対話エージェントや運転支援に関する既存研究について述べる。第 3 章では、運転支援を目的とした合成音声の生成方法に関して示す。第 4 章で制作した合成音声に関するアンケートの分析結果を報告し、第 5 章で本稿をまとめる。

## 2. 関連研究

本章では運転支援を目的としたエージェントに関する研究について述べる。片上らは文献 [7] において、自動車を擬人化したエージェント **HAL Talk Agent** を開発し、ドライバーと自動車の新しい関係を構築することを目指した。擬人化されたエージェントは、自動車の事故診断機能 (OBD) から取得した情報を基に、ドライバーと適切なインタラクションを行う。実験結果から、エージェントが被験者の運転に対する印象を改善することが示された。

また、田中らは文献 [2] において、エージェントの形態が、高齢運転者の注視行動にどのような影響を与えるかを調査した。ここで、エージェントの形態とは、スピーカーからの音声、ディスプレイの映像、ロボット (ヴィストン株式会社の Sota<sup>\*2</sup>) の 3 種類を指している。高齢運転者の主観評価では、ロボットが高齢者にとって最も受け入れやすく、煩わしさが少ないという結果を示した。また、スピーカーからの音声の評価が最も低かった。注視行動において

は、ディスプレイの映像やロボットは、注視点がまとまりやすく、集中した運転を導くことができた。一方、スピーカーからの音声は、注視点が発散しており、高齢者にとっては情報が伝わりにくい手段であることが示された。

これらの研究結果では、映像やロボットを媒介に用いたエージェントが優れた評価を得ている。また、音声の特徴には言及されておらず、合成音声は運転者に与える効果は明らかにされていない。機械的な表現ではなく、感情を込めた表現が可能なエージェントとの関係性は、運転支援において大きな影響を与える可能性がある。そこで、本研究では、感情表現が可能な合成音声を利用したエージェントが、被験者に与える印象を調査する。

## 3. 運転支援を目的とした合成音声の生成

本章では、運転支援を目的とした合成音声の生成方法に関して述べる。まず、生成した合成音声のパターンに関して述べ、次に、合成音声を用いたドライビング・シミュレータの映像に関して述べる。最後にエージェントの発話内容の一覧を示す。

### 3.1 合成音声のパターン

合成音声は感情表現が可能な株式会社エーアイの **AITalk WebAPI**<sup>\*3</sup> を採用した。AITalk では、標準語を発話する 15 種類の話者と、関西風の発話を行う 2 種類の話者が選択可能である<sup>\*4</sup>。また、音量、話速、ピッチ、声の高さの 4 種類の特徴量パラメータ、同様に、喜び、悲しみ、怒りの 3 種類の感情パラメータを調整した合成音声を生成することができる。

本稿では、上記の AITalk を利用して表 1 の 3 パターンの合成音声を生成し比較する。パターン A は、“爽やかで若々しい” が特徴の「のぞみ」を話者とした。感情の表現は用いず、市販のカーナビゲーションの音声のように淡々と発話する。ドライバーに対して冷静で落ち着いた印象を与えることが予想される。パターン B は、上記と同じ「のぞみ」を話者とするが、ドライバーに注意を促す場合には怒りを、運転を褒める場合には喜びの感情を込めて発話する。これにより、運転状況に応じたニュアンスをドライバーに伝える効果が期待できる。パターン C は、“誠実なトーン” が特徴の「せいじ」を話者とした。感情の表現は用いず、パターン A と同様に落ち着いた印象を与えることが期待できる。ここで、音量、和速、ピッチ、声の高さの 4 種類の基本的なパラメータに関しては、共通の値を設定した。

上記の 3 パターンで「もう少し左によりましょう」を発話し、周波数スペクトルのピーク周波数を調べた。パターン A は 242.2Hz、パターン B は 479.1Hz、パターン C は 129.2Hz であった。このことから、作成した合成音声は一

<sup>\*2</sup> <https://www.vstone.co.jp/products/sota/>

<sup>\*3</sup> <https://www.ai-j.jp/cloud/webapi/>

<sup>\*4</sup> 2018 年 1 月現在

表 1 生成した合成音声のパターン

パターン	話者	声の性別	感情	ピーク周波数
A	のぞみ	女性	無	242.2Hz
B	のぞみ	女性	有	479.1Hz
C	せいじ	男性	無	129.2Hz

一般的な話し声の周波数帯とされる、男性 60~260Hz、女性 120~520Hz の範囲内にあることが分かる。また、感情表現を用いたパターン B の周波数が最も高い値を示していることが特徴的である。

### 3.2 ドライビング・シミュレータの映像

前述した合成音声のパターンをドライビング・シミュレータの映像に合わせて再生する。ドライビング・シミュレータの運転映像は、名古屋大学 未来社会創造機構 モビリティ研究 人間・加齢特性研究グループが、エージェント実験で用いているドライビングコースを採用した。このドライビングコースは、株式会社フォーラムエイトが提供している UC-win/Road<sup>\*5</sup> で開発されており、一時停止交差点、歩行者回避、駐車車両回避、一般合流の危険予知などの運転状況で構成されている。図 1 は、一時停止が必要な交差点であり、ドライバーは一時停止の後に、左右確認をした上で、発進することが求められる。図 2 は、道路左に歩行者が存在している状況であり、ドライバーは右方向への迂回行動が求められる。ドライビング・シミュレータの実行時にディスプレイに表示される映像を録画し、映像に合わせて合成音声を再生する。



図 1 一時停止交差点

### 3.3 エージェントの発話内容

表 2 がエージェントが発話する合成音声の一覧である。運転映像は 2 分 17 秒で構成されており、映像中の 14 箇所合成音声による発話を行う。

各発話の目的と、ドライバーの属性を表 3 にまとめる。エージェントの発話目的は、注意喚起と運転称賛に分かれ

\*5 <http://www.forum8.co.jp/index.html>



図 2 歩行者回避

る。注意喚起は、ドライバに交差点や歩行者などの認識を促し、危険運転を指摘する。運転称賛は、ドライバの運転を褒め、安全運転を継続してもらうことを目的としている。14 回の発話のうち、注意喚起が 11 回、運転称賛が 3 回含まれている。

ドライバーの属性は、安全運転と危険運転に分かれる。安全運転は走行速度などに問題がない状況、また、危険運転は速度超過や走行レーンに問題がある状況を表している。14 回の発話のうち、安全運転が 4 回、危険運転が 2 回含まれている。パターン B では、ドライバーの属性に合わせて、感情パラメータを変更する。属性が危険運転の場合は怒り、また、安全運転の場合は喜びのパラメータを上限値の 1 に設定し、感情の抑揚を表現した。また、発話する文字列に句読点をいれるなどして、合成音声为人間的で不自然に聞こえないよう調整した。

## 4. アンケート結果の分析

本章では、被験者 20 人を対象に合成音声に関するアンケートを実施した結果について述べる。アンケートでは、セマンティック・ディファレンシャル法に基づき、上記で述べた 3 パターンの合成音声の特徴を明らかにする。

### 4.1 実験環境

被験者は運転免許を所持した 20 代の女性であり、全員が椙山女学園大学の学生である。本アンケートは予備的な調査であることに注意されたい。実験風景を図 3 に示す。被験者はドライビング・シミュレータの映像が再生されるモニターの前に座り、ヘッドフォンで映像に合わせた合成音声を聞く。スピーカーではなくヘッドフォンを利用することで、より鮮明に音声の区別が可能と考えた。ここで、被験者は先に述べた 3 パターンの合成音声を比較する。このとき、順序効果が生じないように、再生する合成音声のパターンの順番はランダムにした。

表 2 エージェントの発話内容

経過時間	運転状況	発話内容
0:00	停止	今日も安全運転を心がけましょう
0:17	交差点認識	交差点です
0:19	交差点通過	
0:22	路上駐車認識	前方に車が停まっています
0:24	路上駐車車通過	
0:30	交差点認識	
0:32	交差点通過	もう少し速度を落としましょう
0:36	一時停止看板認識	一時停止があります
0:39	一時停止無視	
0:44	交差点認識	交差点です
0:49	交差点通過	左右の確認は大事ですよ
1:00	歩行者認識	歩行者です
1:04	歩行者通過	
1:15	一時停止看板認識	一時停止があります
1:22	右折	丁寧な一時停止ができましたね
1:36	カーブ認識	この先カーブです
1:40	カーブ侵入	
1:43	右寄り運転	もう少し左に寄りましょう
1:50	一時停止看板認識	一時停止があります
1:59	一時停止	
2:01	左折	非常に良い運転です

表 3 エージェントの発話目的とドライバーの属性

経過時間	発話目的	ドライバーの属性
0:00	注意喚起	安全運転
0:17	注意喚起	
0:22	注意喚起	
0:32	注意喚起	危険運転
0:36	注意喚起	
0:44	注意喚起	
0:49	運転称賛	安全運転
1:00	注意喚起	
1:15	注意喚起	
1:22	運転称賛	安全運転
1:36	注意喚起	
1:43	注意喚起	危険運転
1:50	注意喚起	
2:01	運転称賛	安全運転

## 4.2 アンケート結果と考察

アンケートでは、比較する3つの合成音声に関して、セマンティック・デファレンシャル法に基づき、表4に示す12種類の指標の評価を得ることを目的とした。各評価に対し、「全く思わない」、「思わない」、「あまり思わない」、「どちらともいえない」、「やや思う」、「そう思う」、「とても思う」の7段階で評価してもらった。ここでは、上記の7段階の評価を1から7のスコアに換算し、その平均値をレーダーチャートで示す。レーダーチャートは、各合成音声のスコアに加え、3パターンを合わせた全体の平均値を示している。



図 3 実験風景

表 4 アンケートの設問

番号	設問
1	音声に違和感がなかった
2	音声は聞き取りやすかった
3	音声のタイミングは適切であった
4	音声にとっても好感が持てた
5	音声を不快に感じなかった
6	危険な走行場面に音声があっていた
7	安全な走行場面に音声があっていた
8	音声の運転支援は頼りになる
9	音声の運転支援に安心感がある
10	音声の運転支援によって安全に運転ができると思った
11	音声の運転支援を参考にして運転したいと思った
12	自分の両親や祖父母に使って欲しい

### 4.2.1 パターン A

パターン A (女性, 感情無) の結果が図4である。音声の違和感・不快感の値は共に低く、聞き取りやすさ・好感の評価が最も高かった。一般的なカーナビの合成音声に最も近い特徴を持つことが理由と考えられる。自由記述には、「教習所を思い出す、聞き馴染みのある声だった」や「カーナビ音声のような無機質な音声がいい」という意見があったことも、この結果を裏付けている。一方で、「音声の運転支援は頼りになる」と「危険な走行場面に音声があっていた」は、最も評価が低く、注意喚起の効果は高くないことが明らかになった。

### 4.2.2 パターン B

パターン B (女性, 感情有) の結果が図5である。「安全な走行場面に音声があっていた」と「危険な走行場面に音声があっていた」の評価が最も高かった。これは、感情を込めた音声の抑揚が、ドライバーの状況に合っていたことを意味している。また、「音声の運転支援は頼りになる」、「音声の運転支援に安心感がある」の評価も高く、人間的な抑揚のある音声は、被験者の緊張を和らげる効果があることがわかる。一方で、「音声を不快に感じなかった」と「自分の両親や祖父母に使って欲しい」の評価が低い。自由記述では「音声に抑揚があって聞き取りやすかった」や「女



図 4 パターン A (女性, 感情無) の評価

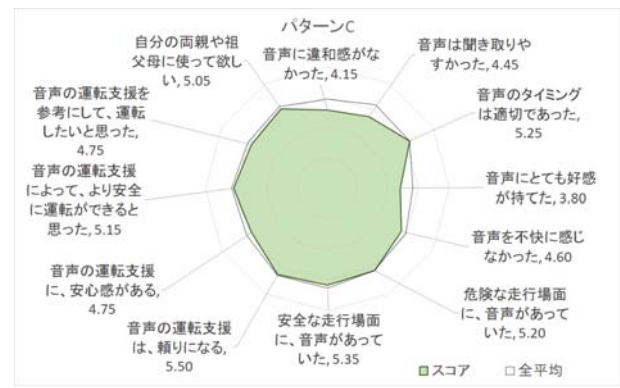


図 6 パターン C (男性, 感情無) の評価

性の高い声が聞きやすい」という肯定的な意見がある一方で、「抑揚が強すぎてイライラする」や「注意のアナウンスに危機感がない」など、否定的な意見も混在していた。被験者の性格や特徴に応じて感情は使い分ける必要があると思われる。

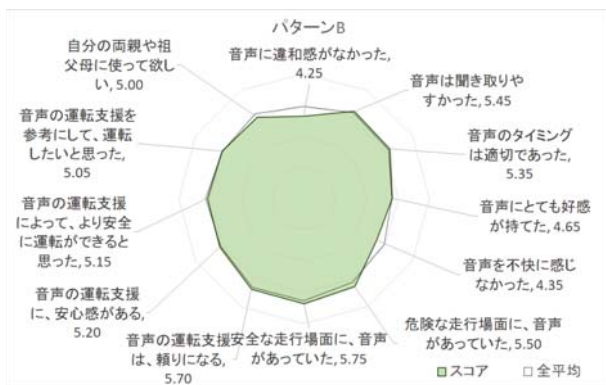


図 5 パターン B (女性, 感情有) の評価

取りやすさ」、「違和感のなさ」、「好感」、「不快感のなさ」、「両親や祖父母に使用して欲しい」など、多くの項目において高く評価された。これは、既存の多くのカーナビゲーションが女性の声を採用していることや、被験者全員が女性であったことが理由に考えられる。パターン B (女性, 感情有) は、安全・危険な場面に最も適した音声であるという評価を得た。一方で、「抑揚が強すぎる」、「声が高すぎる」という否定的な意見が目立った。運転支援という用途において、感情を込めた合成音声は一部の被験者においては有効であるが、強すぎる抑揚は却ってストレスをもたらす可能性があることがわかった。実用化には抑揚のバランスの調整が必須と考えられる。

今後は異なる性別や年齢の被験者で実験し、評価にどのような差が生じるかを明らかにしたい。また、今回は、危険な運転を想定した合成音声を作成したが、長期運転を想定した雑談等の音声を作成することで、同乗者・パートナーとしての効果も検討したい。

## 謝辞

本研究は名古屋大学 未来社会創造機構 モビリティ研究人間・加齢特性研究グループの「エージェントを介した運転支援研究プロジェクト」の支援を受けている。研究グループの関係者に深く感謝する。

## 参考文献

- [1] 下川原 (佐藤) 英理, 篠田遥子, 李 海妍, 高谷智哉, 和田一義, 山口 亨: 高齢者と音声対話ロボットの雑談履歴の解析, 日本ロボット学会誌, Vol. 34, No. 5, pp. 309-315 (2016).
- [2] 田中貴紘, 藤掛和広, 米川 隆, 山岸未沙子, 稲上 誠, 木下史也, 青木宏文, 金森 等: 運転支援エージェントの形態と注視行動の関係分析: 高齢ドライバーの運転行動改善を促すドライバエージェント研究 (人工知能と知識処理), 電子情報通信学会技術研究報告, Vol. 117, No. 72, pp. 13-18 (2017).
- [3] 平井啓之: 音声合成エンジン AITalk を用いた事業展開と最近の取り組みについて, 技術報告 5, (株) エーアイ (2017).
- [4] 徳田恵一: 風雲急を告げる音声合成研究の最新動向, 情報・システムソサイエティ誌, Vol. 21, No. 4, pp. 10-11

### 4.2.3 パターン C

パターン C (男性, 感情無) の結果が図 6 である。「音声に違和感がなかった」の評価が最低であり、自由記述でも「男性の声は怖い」、「男性だと違和感がある」、「女性の方が頭に入ってきやすい」など、否定的な意見が多くを占めた。一般に、男性の声は女性の声に比べて周波数が低く、自動車の走行音などのノイズと周波数が近いので、聞き取りにくいとされる。運転支援に男性の声は向かないことが示唆された。

## 5. まとめ

本研究では、特徴の異なる合成音声を作成し、その特徴量がドライバーにもたらす運転支援効果の違いを明らかにすることを目的とした。目的を達成するため、ドライビング・シミュレータの走行映像に合わせて合成音声再生される仕組みを構築し、被験者 20 名に対してアンケートを実施した。

実験結果から、パターン A (女性, 感情無) が、「聞き

- (2017).
- [5] 宮本敬子, 日根恭子: 合成音声と人声による発話の長さが音声の印象評価へ与える影響, 日本認知心理学会発表論文集, Vol. 2017, p. 79 (2017).
  - [6] 布目光生, 黒田由加, 芦川 平, 森田真弘: 音訳支援システム DaisyRings の開発と音訳コミュニティでの実証—読むことが困難な人にも楽しめるオーディオブック作成—, デジタルプラクティス, Vol. 8, No. 1, pp. 11–20 (2017).
  - [7] 片上大輔, 本郷雅人, 田中貴紘: 自動車とドライバー間の関係性構築のためのエージェント設計, 日本知能情報ファジィ学会ファジィ システム シンポジウム 講演論文集, Vol. 30, pp. 356–357 (2014).