運転負荷の強弱による発話状況の明確化

城 秀一[1]　神沼 充伸[2]　横窪 安奈[1]　ロペズ ギヨーム[1]

概要：本研究では，カーナビゲーション利用時の音声対話に着目し、運転負荷の高低による発話状況の変化について調査した．人間の自由な発話にも対応可能な疑似カーナビと運転負荷の異なる状況を反映したドライビングシミュレータを用いて，運転者に対して運転タスクが課せられた環境下で音声対話システムとの対話タスクを実施し，ドライバの発話の特徴を調べる実験を行った．分析結果から，言語的特徴はほとんど変化がなかったが，対話的特徴は User Response Time(URT)及び Time For Thinking(TFT)について運転負荷が高いときに有意に増加していた．以上から，運転負荷が高い場合，カーナビへの応答が困難になる可能性が高いことが示唆された．

Clarification of the Effect of Driving Load Strength on Speech Situation

HIDEKAZU SHIRO[1]　ATSUNOBU KAMINUMA[2]
ANNA YOKOKUBO[1]　GUILLAUME LOPEZ[1]

1. 序論

昨今，日本では，自動車所有者の 8 割強の車両にカーナビゲーション（以下，カーナビ）が搭載されており，カーナビの普及率は増加傾向にある[1]．カーナビにより目的地までのルートを簡易に把握できる一方，従来のカーナビは運転操作の間にパネルを視覚的に確認しながら手でスイッチ等を操作しなければならず，ドライバは 2 つの操作を同じ時間帯にこなす必要があった．この問題を解決するため，アイズフリーかつハンズフリーな音声対話型カーナビの導入及び研究が進められている．

運転中にも操作できる音声対話型カーナビの開発に向けて，対話コーパスの収集実験[2]や自由対話を可能にするための研究[3]，運転状況に応じた発話内容の変化を調べた研究[4]などが行われている．しかし，これらの研究では発話の仕方が限定されている点や，対話タスクが目的地設定のみに限定されている点などを踏まえると，将来を見越した音声対話型カーナビとしては不十分である．ドライバの自由な発話に対しても入力が可能で，音楽再生や検索など，目的地設定以外のタスクに関しても操作が可能であるカーナビの必要性を感じた．

筆者らは，運転中にも操作できる音声対話型カーナビの開発を目的とし，運転状況や個人に合わせたカーナビの応答を可能にするべく研究を行っている．本稿では，運転状況の一つである，ドライバにかかる運転負荷とドライバの発話内容に着目し，運転負荷の大きさとドライバの自由な発話に対しても操作可能な音声対話型カーナビに対する発話内容との関係について検証する．

2. 関連研究

伊藤らは，運転中または停車中，対話の相手が人間または機械という違いによって，発話形態に差異が生じるのかを調べている[5]．対話を分析した結果，運転操作の有無による音響的特徴には一部違いが見られたが，言語的特徴には差異が見られなかったことが報告されている．

岩本らは，上記の実験をさらに発展させ，運転状況に難易度を 3 段階設けて実験を行った[6]．事前に経路と目的地を暗記させた後に目的地設定タスクの対話課題を課した．対話を分析した結果，心理的負担の違いは，発話の言語的特徴においてほとんど影響を与えず，音響的特徴に若干の影響を与えるであろうということがわかった．

音声対話型カーナビに必要な要素は，人間の自然で自由な発話に対しても認識できるようにすることであると考える．一方で，ユーザの運転負荷を考慮すると，運転状況が発話の内容や応答タイミングに及ぼす影響が拭えず，運転状況ごとに音声対話型カーナビが受諾すべき言語的特徴や対話のデザインが変わるかもしれない．そこで本稿では，運転負荷が異なることで，ユーザの自由な発話に対しても操作可能な音声対話型カーナビとの発話状況に差異があるのかを検証する．このため，運転負荷の高い場合と低い場合について，言語的特徴では品詞による比較，対話的特徴では反応時間やタスク達成時間で比較を行った．

1 青山学院大学
Aoyama Gakuin University
2 日産自動車株式会社
NISSAN MOTOR CO., LTD.

ⓒ2019 Information Processing Society of Japan

3. 実験環境の構築と実験条件

3.1 ドライビングシミュレータ

運転操作を実現するために株式会社フォーラムエイトが販売する UC-win/Road ドライブ・シミュレータ[7]（以下，DS）（図 1）を用いた．DS は，前方および左右側面（全 3 面）のディスプレイへの表示，バックミラー位置に設置された後方画面，フォースフィードバック機構が搭載されたステアリング，ターンシグナル，足踏み式のアクセル・ブレーキ，オートマチックトランスミッション操作を模擬したシフトで構成されている．走行コースは 8 の字型で，直線と緩やかなカーブのみの 2 車線で構成されている．天気は晴れ，信号機や合流地点は設置しない．ドライバが運転する車両は，右ハンドルの AT 車と統一した．

運転負荷を統制するために，ドライバの運転する車両の前方に 1 台の先行車を配置する．この車両は，車線（右・左）と速度（40km/h・80km/h）を実験者が変更できる．

図 1　ドライビングシミュレータを用いた運転の様子

3.2 疑似カーナビゲーションシステム

カーナビとの自然で自由な対話を実現するために，Wizard of OZ 方式[8]（以下 WOZ 方式）による疑似カーナビを作成した．これはドライバが音声であたかもカーナビを操作していると擬似的に感じられるシステムである．

WOZ ではオペレータが操作画面にて特定の応答文が割振られたスイッチを押すことによって，ドライバ側に設置してあるカーナビ画面の表示と，応答文の音声呈示（イヤホンから呈示）が可能である．例えば，ドライバから「鉄道博物館を目的地に設定する」との音声入力があった場合，オペレータは「鉄道博物館を目的地に設定しました」と応答文を２、３回のスイッチ操作で入力すると，システムから当該音声が読み上げられ、同時にドライバの左前方に設置されたカーナビ画面を変更される（図 2）．オペレータによる操作時間は，概ね 3 秒以内である．

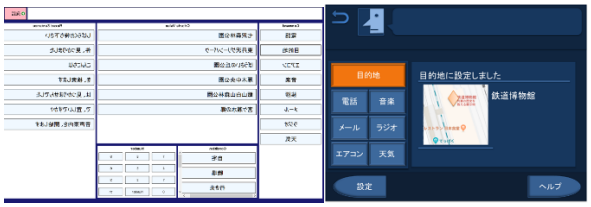

図 2　WOZ 操作画面（左）とカーナビ画面（右）

3.3 実験環境の構築

データの収録には，8 トラック同時収録可能なズーム社製のマルチトラックレコーダ ZOOM R24[9]（以下，MTR）を用いた．1 から 6 トラックまでを使用し，それぞれ「オペレータ音声」「ドライバ音声」「WOZ 音声」「課題提示ガイド音声」「楽曲再生用音源」「ドライバ心電図」として収録した．また，録画用ビデオカメラを 3 台用いて，「ドライバ顔表情」「ドライバ全体」「DS の中央ディスプレイ」の 3 視点が録画できるように設置した．上述の録画に加え，生体信号（心電・脳血流・視線）の収録も同時に実施した．実験環境のレイアウトは，図 3 の通りである．

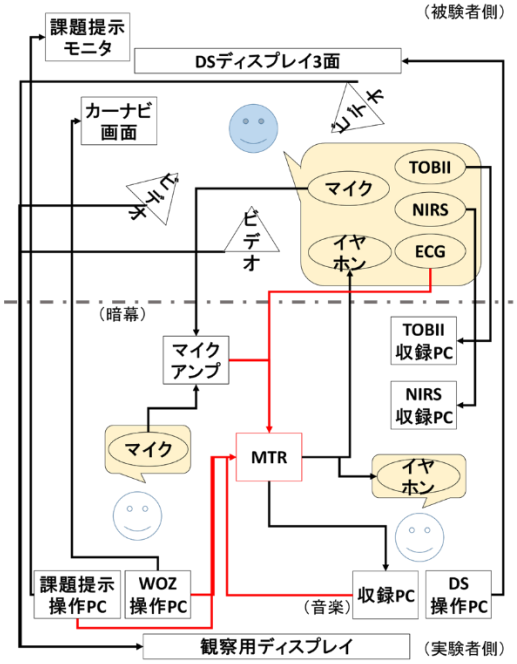

図 3　実験環境のレイアウト

3.4 実験条件

運転負荷が異なる運転状況の中で，ドライバには音声で操作するカーナビと対話を行い，課題を進めることを求めた．そのときの対話音声等から，運転負荷の違いによる発話特徴について分析を行った．本実験のドライバは，本学に在籍する 20 代の男女 13 名と，一般のドライバ 30 代から 60 代の男女 20 名を対象とした．ドライバは，普通自動車運転免許証を所持していることを条件とした．

実験条件および実験は，日産自動車株式会社の人を扱う実験倫理委員会の審査を経て実施された．

3.4.1 対話タスク

音声のみで操作するカーナビとの対話を行うために，本実験では 18 個の課題を設定した．これらの課題は一般的にカーナビで操作するタスクである目的地設定・検索，ラジオ・音楽再生，メール送受信，電話，音量調節,および地図の縮尺調節を含めた設定変更であり，現実のカーナビの使用状況を想定し，事前に用意したストーリーに沿うように設定した（表 1）．ドライバの負担を考慮し，全 18 課題を

ⓒ2019 Information Processing Society of Japan

前半 10 課題と後半 8 課題に分け，さらに各パート内で 1 回休憩を挟むことで，計 3 回の休憩時間を設けた．課題の提示は，詳細な説明（課題提示ガイド音声）をドライバのイヤホンから再生し，簡易的な説明をドライバ左情報に設置した課題提示モニタにて表示した．画面の呈示例を図 4 に示す．

表 1　対話タスク一覧

課題番号	機能	シナリオ	実施区分
1	検索	あなたは今，渋谷で車を運転しています． これから友達を車に乗せて横浜か大宮に遊びに行く予定です． まず最初に，横浜と大宮の天気を調べましょう． 天気の良い方へ向かいます．	前半パート1
2	目的地設定	天気を調べたあなたは，鉄道博物館で観光することにしました． 鉄道博物館を目的地にセットしましょう．	
3	メール	あなたは渋谷駅で友達を車に乗せる予定です． メールボックスを開いて彼からメールが来ていたら読みましょう．	
4	メール	友達から待ち合わせ場所の変更を提案されました．このメールに返信をしましょう． なお返信の際は，ナビの返信画面に表示されている定型文のうち一つを選んで返信してください．	
5	検索 目的地設定	待ち合わせ時間までに余裕があるため，あなたはコンビニでお茶を買うことにしました． コンビニを検索し，出てきた候補のうち行きやすい地点を経由地に設定しましょう． あなたはセブンイレブンのポイントカードを愛用しています．	
6	電話	そろそろ駐車場に到着しそうです．さきほど待ち合わせのメールをした友達に電話をしましょう．	
7	ラジオ再生	無事友達を車に乗せることが出来ました． ラジオで人気の歌番組が今から始まるので，ラジオを起動して歌番組が流れているチャンネルに変えてみましょう．	前半パート2
8	音楽再生	ラジオを聞いた友達が別のコブクロの曲を流してほしいと言っています． 「さくら」を流しましょう．	
9	検索 目的地設定	お腹が空いたので昼ごはんを食べたいと思います．あなたは今走っている国道沿いに美味しいうなぎ屋があることを思い出しました． このお店を調べて経由地に設定しましょう．	
10	電話	これから行くうなぎ屋は人気があるため，事前に予約をした方がよさそうです． うなぎ屋に電話をしましょう．	
11	音楽再生	うなぎを食べて車に戻りました． 欅坂46の太陽は見上げる人を選ばないを流しましょう．	後半パート1
12	設定変更	音楽を流しましたが，どうやら音が小さいようです． あなたの聞こえやすいように音量を調節しましょう．	
13	設定変更	鉄道博物館周辺に来ました． 博物館周辺の地図が細かくて見辛いため，駐車場の位置が分かるように地図の縮尺を調節しましょう．	
14	音楽再生	友達が欅坂46の続きを聞きたいと言っています． 楽曲を流しましょう．	後半パート2
15	目的地設定	鉄道博物館を楽しんだあなた達は，友達を大宮駅に送って帰宅することにしました． 大宮駅を目的地に設定しましょう．	
16	目的地設定	大宮駅に向かっている途中ですが，まだ少し時間があるので，お洒落な御朱印がもらえると噂の氷川神社に行くことにしました． ここを新たな経由地に設定しましょう．	
17	検索 目的地設定	友達に急用が出来てしまったため，駅や神社には行かず，友達の家まで直接送ることにしました． 友達の家の住所を電話帳から調べ，目的地に設定しましょう．	
18	電話	友達を駅に送り届けました． 先ほど着信あった外国人の先生に折り返しの電話をしてから帰宅しようと思います． 着信履歴から先生を探して電話を掛けましょう．	

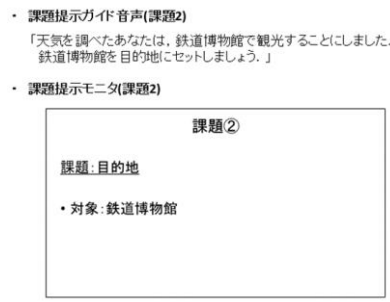

図 4　課題の呈示例

3.4.2 運転タスク

運転負荷の違いを実現するために，運転負荷が低いパートと高いパートを対話タスクの前半後半にて分けた．課題の順序効果を考慮し，ドライバによって「前半: 低，後半: 高」「前半: 高，後半: 低」となるように入れ替えた．

運転負荷の低いパートは，制限速度 60km/h にて走行するだけで，先行車は存在しない．運転負荷の高いパートは，先行車に追従する．先行車はオペレータの操作によりランダムで，左右の車線変更と急加速（80km/h）と急減速（40km/h）を行う．追従間隔は道路中央の白点線 3 つ分を目安とし，ドライバには安全運転を最優先として走行することを求めた．

4. 分析手法

本実験で得られた音声対話型カーナビとの対話音声データおよび当該音声を書き起こしたテキストデータを用いて，運転負荷の高低による言語的特徴と対話的特徴について分析する．

言語的特徴を調べるため，「Praat」[10]にて対話内容の書き起こしと対話開始・終了時間のマーキングを行った．書き起こされた対話内容のうち，ドライバの発話については「MeCab」[11]にて形態素解析を行い，単語毎に品詞を求めた．この発話内の品詞（10 種: 名詞，動詞，形容詞，副詞，助詞，接続詞，助動詞，連体詞，感動詞，その他）の出現回数を発話ごとに数え上げた．また，ドライバの第 1 発話目は課題を解決するための意図を発話している場合が多いため，第 1 発話目の各品詞の出現数を別途数え上げた．

次に，対話的特徴を調べるため，参考文献[12]に記載の以下のパラメータを測定，計算した（図 5）．

- システム発話～ユーザ応答までの時間（URT）
- 課題提示直後～最初のユーザ発話までの時間（TFT）
- 課題提示後の最初のユーザ発話～課題達成を示す最後のシステム応答までの時間（TTT）
- タスク達成時間（TFT+TTT）

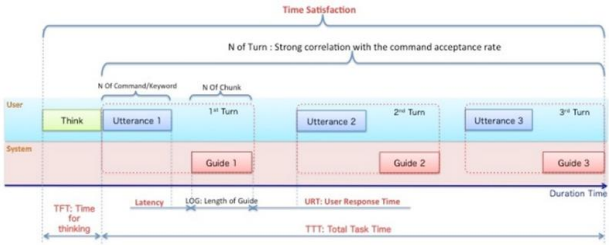

図 5　対話的特徴の定義（[12]から引用）

これら評価尺度について，対話タスクの 18 課題の中で機能分類の「目的地設定」「メール」「検索」「音楽再生」「電話」にあたる 8 課題（タスク 4 種類×運転負荷条件 2 種類）の分析を行った．

5. 分析結果と考察

5.1 言語的特徴

運転負荷の違いによる第 1 発話目の各品詞の出現数に対して，Wilcoxon の順位和検定を行った．各課題内では，運転負荷の違いによる有意な差はないことが確認できた（$P<0.05$）．

ⓒ2019 Information Processing Society of Japan

次に，課題内の運転負荷低と高で全被験者の第 1 発話目の各品詞の出現数の平均と分散をそれぞれ算出した．平均と分散を指標として全課題に対して，Wilcoxon の符号付順位和検定を行った．運転負荷低に比べ，高は，名詞の平均には有意な差は見られなかったが分散に有意な増加（図 6），動詞の平均と分散に有意な減少（図 7）が確認できた（P<0.05）. 図の縦軸は平均出現数, 横軸は課題番号を示す. 棒グラフのひげは標準偏差を示す。棒グラフの値は,数え上げた名詞および動詞の平均出現数を示す.

しかし，表 2 の示すように，運転負荷を変えても出現回数の差は有意に変わらず，分析対象として有効な値とは考えにくい．以上から，運転負荷の違いによる言語的特徴の有意な変化は検出されないことが分かった.

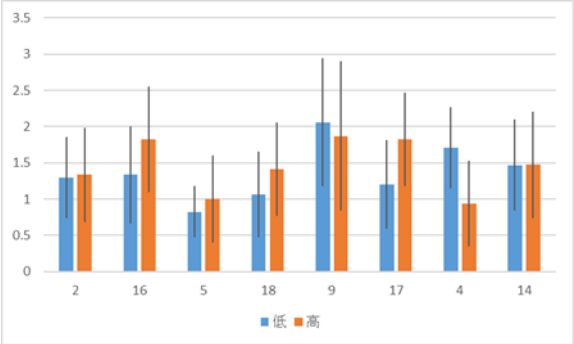

図 6　言語的特徴（名詞）における運転負荷高低の関係

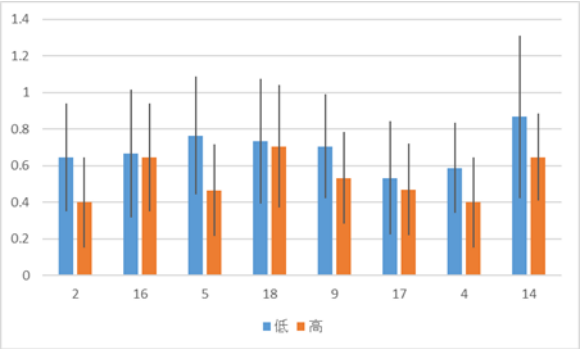

図 7　言語的特徴（動詞）における運転負荷高低の関係

表 2　言語的特徴の名詞（左）と動詞（右）における運転負荷高低の関係

低	平均(回)	1.4	低	平均(回)	0.7
	分散	1.6		分散	0.4
高	平均(回)	1.5	高	平均(回)	0.5
	分散	2.1		分散	0.3

5.2　対話的特徴

運転負荷の違いによる URT と TFT に対して，t 検定を行った．運転負荷低に比べ，高は，課題 9（検索・目的地設定）では TFT に有意な増加が確認できた（P<0.05）.

次に, 課題内の運転負荷低と高で全被験者の URT と TTT と TTT+TFT と TFT の平均と分散をそれぞれ算出した．平均と分散を指標として全課題に対して，t 検定を行った．運転負荷低に比べ，高は，URT の平均と分散に有意な増加（図 8），TFT の平均に有意な増加（図 9）が確認できた（P<0.05）．図の縦軸は平均時間（秒），横軸は課題番号を示す．棒グラフのひげは標準偏差を示す。棒グラフの値は, 計測した URT および TFT の平均秒数を示す.

表 3 の示すように，URT の平均と分散を見ると，運転負荷が高いとシステムへの入力に 1 秒多く時間がかかっていることが分かった．さらに，TFT にも有意な差があり，課題の解決方法を考えるのにも 0.6 秒多く時間がかかっていることが分かった.

以上から，運転負荷の違いによる対話的特徴の変化は，URT と TFT に有意な差があることが分かった. 運転負荷が高い場合は，システムへの入力に多く時間がかかってしまっていると考えられる.

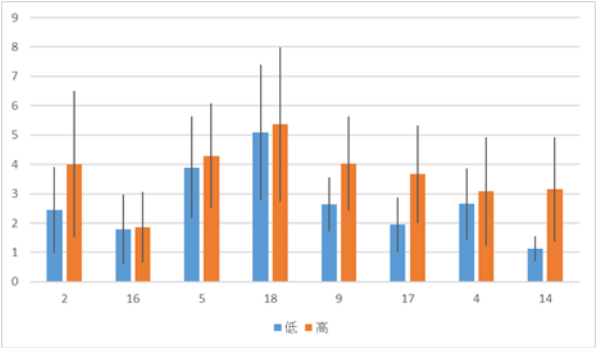

図 8　対話的特徴（URT）における運転負荷高低の関係

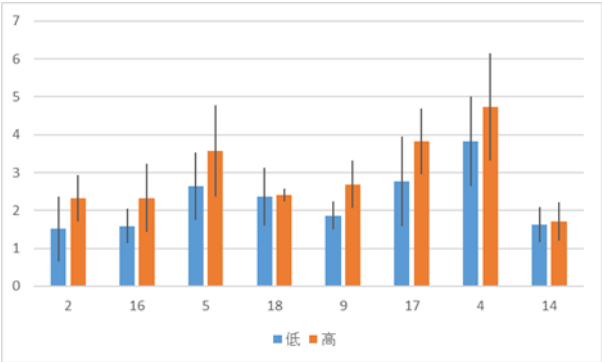

図 9　対話的特徴（TFT）における運転負荷高低の関係

表 3　対話的特徴（URT・TFT）における運転負荷高低の関係

		URT	TFT
低	平均(秒)	2.7	2.3
	分散	7.6	2.7
高	平均(秒)	3.7	2.9
	分散	14.8	3.0

5.3　考察

先行車の有無による運転負荷の違いによって，言語的特徴の有意な変化は検出されなかった。この結果より、ドライバによる運転中の発話は，予め記憶領域に発話する文章を保存し，タイミングを待って当該記憶された内容そのものを発話しているかのように見える．この記憶領域に保存された文章は，運転負荷に応じてその内容がドライバによって修正されていないことから，ドライバは運転操作と音声発話の 2 重課題を実施する際には，音声発話の内容を短

ⓒ2019 Information Processing Society of Japan

時間で修正するだけのリソースを有していない可能性が示唆される．

一方で，対話的特徴の変化は有意に存在していた．これらの結果は，運転負荷が高いときは，低いときに比べて，発話した文章内容は同じであるにも関わらず，システムからのガイダンスを聞いた後，発話を開始するまでの時間は多くかかっているということを示している．この原因として 2 つの考察が考えられる．一つは，運転負荷が高いときには，ドライバが運転操作に集中してしまい，運転負荷が低い場合と比較して，システムの応答内容を理解するまでの時間が長くなってしまうこと，もう一つは，予め記憶した文章を思い出すまでに時間がかかってしまうことである．これらの知見は音声対話型カーナビの利便性を高める要素として重要であり，今後は運転状況を把握し，運転操作が難しくなると判断した場合に応答を一時中断するなどの機能を搭載するなどの工学的発展が期待できる．

6. 結論と今後の展望

分析結果から，運転負荷の違いによって，言語的特徴の変化はほとんどなく，対話的特徴の変化があることが明らかになった．運転負荷が低い時と比較すると，運転負荷が高い時は，発話内容が同一であっても発話までの時間を多く要することが示唆された．これは，運転負荷が高い，すなわち運転時の判断が増加することで，その判断に集中してしまい，カーナビの応答内容を理解することが困難になっていると考えられる．そのため，音声対話型カーナビの利便性を高める要素として，運転状況を把握し，運転操作が難しくなると判断した場合に応答を一時中断するなどの機能を搭載するべきであると言える．

今後の展望として，より精度の高い結果を得るために，本研究で限定した 8 課題に留まらず，18 課題すべてを分析対象とすることが挙げられる．また，各課題の難易度を考慮するために，課題達成者だけではなく，未達成者についてもどのような変化が起こっているのか分析する必要がある．誤認識をした際のユーザの返答を分析することで，より自由な発話にも対応できるようになる．課題内容については，複雑にしすぎてしまい，課題達成者が少なくなってしまったので，単純でわかりやすい課題を再検討する必要がある．

本研究では対話コーパスに着目したが，生体情報（心電図・脳血流・視線）・運転情報（速度・アクセル量・ブレーキ量・ステアリング角度）・個人特徴量の影響も考えられる．これらの情報と発話内容を関連付けることで，運転者に合わせた音声対話型カーナビの実現が可能になると考える．

謝辞

実験作業に協力していただいた，青山学院大学 Lopez 研究室の内田昂氏に感謝いたします．

参考文献

[1] マイボイスコム株式会社: “カーナビの利用に関するアンケート調査（第 9 回）”, https://myel.myvoice.jp/products/detail.php?product_id=24504, （参照 2019/5）

[2] 松原茂樹, 河口信夫, 外山勝彦, 武田一哉: “音声対話コーパスの収集と利用　より豊かな車内音声対話システムを目指して”, 人工知能学会誌 17 巻 3 号, pp. 279-284 (2002).

[3] 藤本拓, 原隆浩, 西尾章治郎: “自然な発話により操作可能なカーナビゲーションシステムの開発”, 電子情報通信学会論文誌, Vol.J96-D No.11, pp. 2815-2824 (2013).

[4] 岩本善行, 伊藤敏彦, 甲斐充彦, 小西達裕, 伊東幸宏: “同時処理タスクの難易度の変化における言語能力への影響”, 自然言語処理, pp. 149-19 (2002).

[5] 伊藤敏彦, 甲斐充彦, 岩本善行, 水谷誠, 由浅裕規, 小西達裕, 伊東幸宏: “目的地設定タスクにおける対話状況の違いによる言語・音響的特徴の比較”, 情報処理学会論文誌, Vol.43-No.7, pp. 2118-2129 (2002).

[6] 岩本善行, 伊藤敏彦, 甲斐充彦, 小西達裕, 伊東幸宏: “同時処理タスクの難易度の変化における言語能力への影響”, 自然言語処理, pp. 149-19 (2002).

[7] 株式会社 FORUM8: “UC-win/Road”, http://www.forum8.co.jp/product/ucwin/road/ucwin-road-1.htm, （参照 2018/12）

[8] Bernsen, Dybkjær, et al., "Wizard of OZ prototyping: How and when." Proc. CCI Working Papers Cognit. Sci./HCI, 1994.

[9] 株式会社 ZOOM: “R24”, https://www.zoom.co.jp/ja/products/multi-track-recorder/r24-recorder-interface-controller-sampler, （参照 2018/12）

[10] Paul Boersma, David Weenink: “Praat”, http://www.fon.hum.uva.nl/praat/, （参照 2018/7）

[11] 京都大学情報学研究科-日本電信電話株式会社コミュニケーション科学基礎研究所: “MeCab”, http://taku910.github.io/mecab/, （参照 2018/11）

[12] 神沼充伸, 網本侑理, “ドライバのストレスレベルに着目した音声対話システムの評価法”, JSAE2018 春季大会

ⓒ2019 Information Processing Society of Japan