

カーナビ音声操作の発話タイミングミス軽減

Mitigating Difficulties of an Utterance Timing for The Speech Recognition in Car Navigation Systems

外山 聡一, 川添 佳洋, 小林 載, 藤田 育雄
Soichi Toyama, Yoshihiro Kawazoe, Hajime Kobayashi, Ikuo Fujita

金子 忠靖, 大杉 淳, 塩田 岳彦
Tadayasu Kaneko, Jun Oosugi, Takehiko Shioda

要 旨 カーナビ音声操作機能の機能向上を目的にユーザビリティテストを実施し、認識阻害要因の調査を行った。その結果、発話タイミングの取りづらさが一つの大きな要因であると判明し、対策を検討した。カーナビ音声操作のプロトタイプツールを作成し、対策方法考案/プロトタイプ作成/少人数被験者へのテストを繰り返して、対策手法の検討・絞り込みを行った。その結果、画面の改良およびユーザ主導での発話タイミング切り替えを効果のある対策手法として選択した。この対策手法のプロトタイプに対して、一般被験者へのユーザビリティテストを行い、発話タイミングミス軽減効果があることを確認した。

Summary In order to improve the ease of use of the voice operation function in our car navigation, we performed a usability test of the function and investigated the factors involved in recognition obstruction. We found that it was difficult for the user to utter a voice command at the correct timing. As a result of examining possible improvement methods, we confirmed that it was effective to make a combination of improvements in the display and the operation sequence together.

キーワード : カーナビ, 音声認識, 発話タイミング

1. まえがき

現在販売されている多くのカーナビには音声認識機能が搭載されている。これにより、走行時でも音声を使ってハンズフリー・アイズフリーでのカーナビ操作が可能となっている。カーナビに搭載されている音声認識エンジンの認識性能は高く、高速道路を走行しているような劣悪な雑音環境下でも90%前後の高い認識率が得られるにも関わらず、「うまく認識してくれないので使わない」といったユーザの声が少なからずあるのも事実である。

今回我々はまず、音声操作に慣れていない一般被験者へのユーザビリティテスト⁽¹⁾を実施し、被験者が実際にカーナビの音声認識機能を使用している様子を観察した。そして、一般のユーザが音声操作を使い

こなすことができるのか、使いこなせないとしたらその要因はなにかを検証した。その結果、被験者が音声での操作を思うようにできない、という場面が見られた。その要因として、正しいタイミングで発話できないという発話タイミングミスの問題と、「何と発話して良いかわからない」という2つの大きな問題が存在することがわかった。この音声操作ユーザビリティテストに関しては第2章に述べる。

ユーザビリティテストで抽出した2つの大きな認識阻害要因のうち、今回は発話タイミングミスに注目し、その軽減法を検討した。多くのカーナビは、音声操作を起動するための発話ボタンが、音声操作リモコンやステアリングなどに用意されている。音声操作のシーケンスとしては、発話ボタン押下(U)→ガイド

ンス音声再生 (N) → ビープ音再生 (N) → 発話 (U) → 認識・ナビ作動 (N) となっている。U はユーザの行動、N はカーナビの動作を示している。また、音声操作にはタイムアウト時間が設定されており、ビープ音再生後数秒間音声入力が検出されないと、タイムアウトし、再び発話ボタンが押されるまで音声入力を受け付けられない待機状態となる。よって、カーナビがビープ音を再生した後、数秒間が発話可能な状態である。ビープ音再生前やタイムアウト後はカーナビが発話を受理できない状態であるが、この間にユーザが発話してしまう問題が発話タイミングミスである。前者はユーザの発話タイミングが早すぎるミス、後者は遅すぎるミスである。さらに、発話可能なタイミングで発話したとしても、発話開始が遅いと、カーナビは周囲の雑音を音声として誤検出してしまい、誤動作を引き起こすといった問題もある。こちらも発話タイミングが遅すぎるミスに含まれる。この2つの問題の内、発話タイミングが早すぎるミスへの対応としては、バージン技術の研究^{(2),(3)}がなされている。これは、音響エコーキャンセラを利用し、マイク入力からガイダンス音声やビープ音を除去する、というものであり、有効である。しかし、特に音声操作に不慣れなユーザは、発話タイミングが遅すぎるミスを起こしやすく、バージン技術だけでは発話タイミングミスの対策として十分ではない。

我々は、音声操作の便利さを体験し、以降活用してもらうためにも、初めて音声操作を利用するユーザが使えるようにすることを第1の目標として検討を進めた。音声操作はハンズフリー・アイズフリーが特徴ではあるが、初めて音声操作を使う場面では、車を停車して画面を見ながら操作するであろうと考えた。そこで、音声操作時の画面表示を有効に活用するという観点で改良方法を検討した。改良案のプロトタイプ作成と少人数被験者へのテストの繰り返しによる最終案絞り込みまでの検討内容を3章に、最終案の効果確認実験の内容を4章に述べる。

2. 現行システムの音声操作ユーザビリティテスト

2.1 テストの概要

上述のように、現在カーナビで実現されている音声操作機能を、音声操作に不慣れな一般ユーザがどのくらい正しく使用することができるのかを調査するため、パイオニア製2003年モデルカーナビを用いてユーザビリティテストを実施した。

被験者は男女各4名の計8名。30～40代のカーナビユーザで音声操作初心者である。被験者に16の課題(タスク)を与え、実行してもらった。16タスクのうち、停車中のタスクは10、走行中のタスクは6(うち3つは高速道路走行)である。実験に使用したタスクを表1に示す。

タスクの構成は、自宅から目的地へ車で行き、用事を済ませて帰ってくる際にカーナビを音声で操作する、という設定になっている。停車中に目的地を設定し、走行中にルートの変更・周辺施設の検索・渋滞情報の確認・地図表示方法の変更などを行う、という内容である。なお、タスク13～16はタスク1～4と同じ種類のタスクとなっている。これは、音声操作を使いこなすことによる学習効果が現れるかどうかを検証するためである。

テストはカーナビを設置した実車で行った。被験者は運転席で音声操作を行う。走行中タスクでは被験者は実際に車を運転した。テストの際には進行係が同乗し、被験者にタスク実行に関する指示を与えた。また進行係は、被験者が操作に行き詰まった場合にアドバイスを与え、被験者全員が全てのタスクを実行できるように導いた。

テストの様子はビデオに撮影し、これを解析して結果を集計した。

表1 音声操作ユーザビリティテストタスク一覧表

タスク番号	走行状態	タスク内容
1	停車	住所での目的地設定
2	停車	電話番号での経由地設定
3	停車	ルートの消去
4	停車	施設名での目的地設定
5	一般道走行	ルートの変更
6	一般道走行	周辺施設の検索
7	一般道走行	地図表示の変更
8	停車	施設名での目的地設定
9	高速道走行	渋滞情報の確認
10	高速道走行	地図表示の変更
11	停車	自宅を目的地に設定
12	高速道走行	周辺施設の検索
13	停車	住所での目的地設定
14	停車	電話番号での経由地設定
15	停車	ルートの消去
16	停車	施設名での目的地設定

2.2 テスト結果

以下にテストの解析結果を述べる。表2にタスク達成率の集計結果を示す。

タスク達成率は、進行係がアドバイスを与えず、被験者が自力で完了できたタスクの割合を指す。全タ

スク平均では約 65% で、半数強のタスクは被験者が自力で完了できた。停車時タスクの平均タスク達成率は走行時タスク平均の約倍で、停車時タスクの達成率が高いことが確認できた。これは、停車時には画面表示を参照できること、走行時には運転しなければならず音声操作に集中できないこと、走行時の雑音の増加による認識性能の低下などが考えられる。

次に表 3 に、前半タスクと後半タスクのタスク達成率を比較した結果を示す。

タスク 1～4 とタスク 13～16 は同じ種類のタスクであるが、実験後半に行ったタスク 13～16 の方が平均タスク達成率が高くなっている。これは、本実験により音声操作を経験したことによる学習効果が現れているものと思われる。

次に、音声認識率を集計した結果を表 4 に示す。ここでは、2 種類の認識率を定義し、集計した。"正しい発話の認識率" と "目的認識率" である。

正しい発話の認識率は、被験者の発話のうち、システムが認識可能な語彙内の言葉を、システムが認識可能なタイミングで発話したもののみを集計して求めた認識率である。発話内容と認識結果が完全に一致した場合のみを正解としてカウントしている。音声認識エンジンの認識性能を測る指標となる。

目的認識率は、被験者がシステムに向かって発話したもの全てを集計して求めた認識率である。システムの認識できない語彙外発話や認識できないタイミングでの発話もカウントしている。ユーザが実際に感じる認識率に対応する。

正しい発話の認識率は、全平均・停車時平均・走行時平均ともに認識率が高く、システムの認識性能が高いことを示している。ただし、停車時と走行時では走行時の認識率が若干低下していることがわかる。これに対し目的認識率は、50% 前後の低い値となっている。走行時は停車時に比べより低い値となっている。

以上の結果から、システムの音声認識性能は高いが、実際に使ってみるとその性能を十分体感できない、ということが確認できた。

続いて目的認識率劣化を引き起こす認識障害要因について集計した結果を表 5 に示す。

認識障害要因としては、被験者がシステムの語彙外の発話をしてしまうことが最も大きな割合となった。次いで、システムがユーザの発話を受け付けられないタイミングで正しく発話できない、という発話タイミングミスによる誤認識が大きな割合となった。こち

らも目的認識率を向上させる上で重要な項目である。さらに、音声操作開始をシステムに伝える、発話ボタンの操作ミスがあった。その他は、音声認識エンジン性能による要因である。

語彙外発話対策は最も大きな課題であるが、これに関しては別途検討することとした。今回はもう一つの大きな要因である発話タイミングミスを軽減する方法に関して検討した。

表 2 タスク達成率集計結果

タスク集計内容	タスク達成率(%)
全平均	64.8
停車時平均	80.0
走行時平均	39.6

表 3 前半タスクと後半タスクのタスク達成率比較

タスク集計内容	タスク達成率(%)
タスク 1～4 平均	65.6
タスク 13～16 平均	87.5

表 4 音声認識率集計結果

認識率集計内容		認識率(%)
正しい発話の認識率	全平均	88.5
	停車時平均	89.8
	走行時平均	86.7
目的認識率	全平均	44.6
	停車時平均	50.3
	走行時平均	37.7

表 5 認識障害要因の集計結果

要因	割合(%)
語彙外発話	67.3
タイミングミス	10.5
語彙外タイミングミス双方	4.5
発話ボタン操作ミス	3.8
その他	13.9

3. 発話タイミングミス軽減方法の検討

3.1 画面表示の改善

現在のカーナビの、音声操作シーケンスフローを図 1 に、画面表示を図 2 に示す。

従来方法では発話ボタンが押されると、キャラクタおよびシステムが認識できる言葉のリストが表示される(図 2 (1))。発話可能な状態になるとキャラクタ

の表示が変わる(図2(2)のでマイクを差し出し、首を傾げる)。ビープ音とこのキャラクタ表示の変化によって、ユーザに発話可能状態への切り替わりを通知しているが、2章で述べたとおり十分理解されていない。そこで、画面表示の改良を試みた。

最初に、実験用のカーナビ音声操作プロトタイプングツールを作成した。そして、3種類の表示改良案のプロトタイプを作成し、少人数の被験者に対してテストを行った。多数の被験者へのテストを頻繁に行うことは、コストの面から困難である。今回我々は、改良案作成と少人数の被験者へのテストを繰り返して対策手法を絞り込んだ。そして、絞り込んだ手法に対して、人数を増やして被験者テストを行い、効果を確認するという手法をとった。

表示改良案は以下である。表示改良案1：従来方法

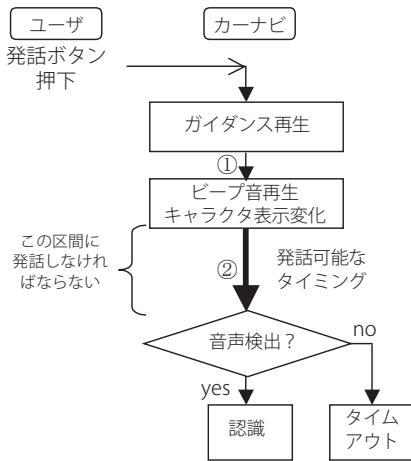
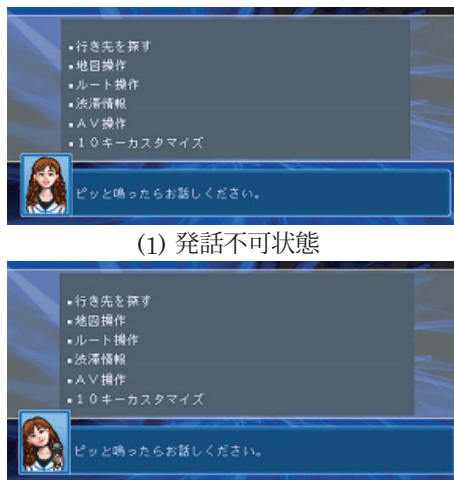


図1 カーナビ音声操作シーケンス



(1) 発話不可状態

(2) 発話不可状態

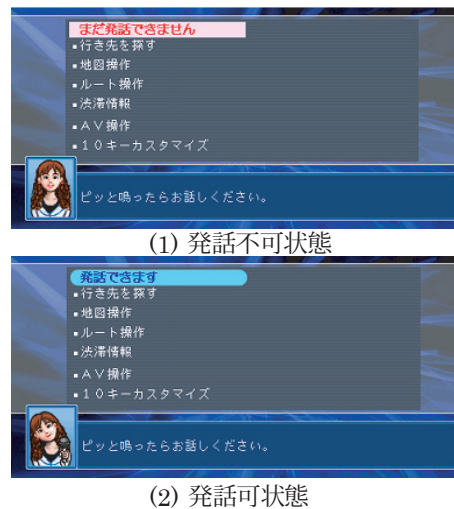
図2 従来方法の画面表示

に加え、リストの横に大きく発話タイミングを通知する発話タイミング通知文字表示を行う。表示改良案2：従来方法に加え、リストと同列に発話タイミング通知文字表示を行う。表示改良案3：表示改良案2を行った上でさらに文字背景の色を変更するようにする。表示改良案2を図3に示す。

この3つの表示改良案を5人の被験者に対してテストを行い、評価した。各被験者に対して、従来方法と3つの表示改良案のうちの1つの案をテストした。被験者一人当たりのタスク数は3~7である。発話可能状態の回数に対する発話タイミングミスを起こさなかった回数の割合(成功率)を、従来方法と表示改良案とで比較したが、差が見られなかった。被験者にインタビューを行ったが、多くの被験者が表示改良案と従来方法の表示の違いに気づいていなかった。その理由は、ガイダンス音声再生されてビープ音が鳴るまでの間(1~2秒)に、リスト表示を参照して何を発話したらよいかを懸命に探していて、リスト表示以外に注意をはらう余裕がないためであることがわかった。そこで、リストをじっくり見られるようシーケンスの改良も同時に行うこととした。

3.2 画面表示およびシーケンスの改善

リストをじっくり見られるようにする対策として以下の3つのシーケンス改良案を試した。シーケンス改良案1：リスト表示後の認識スタートは再度発話ボタンによりユーザ主導で行う。シーケンス改良案2：リスト表示画面を止める"ストップ"コマンドを設ける。シーケンス改良案3：表示するリストを少なくす



(1) 発話不可状態

(2) 発話可状態

図3 表示改良案2ーリストと同列に タイミング通知文字表示

る。シーケンス改良案1のフローを図4に示す。

シーケンス改良案2では,"ストップ"コマンドによりリストを表示した状態で一旦停止し, 発話ボタンにより再開する。シーケンス改良案3では, 短時間での視認性向上を狙って, リストに表示するコマンドの行数を少なくしている。リスト表示を少なくしたため

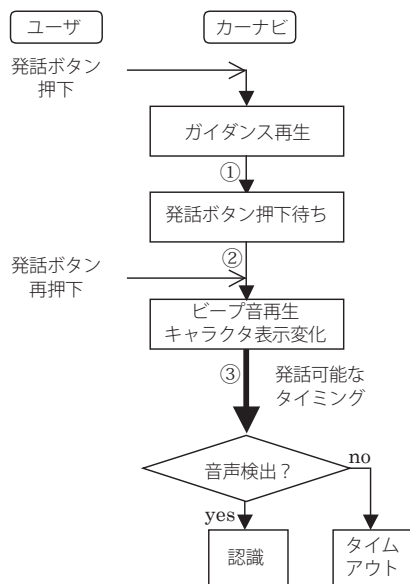
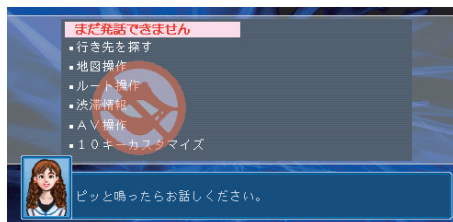
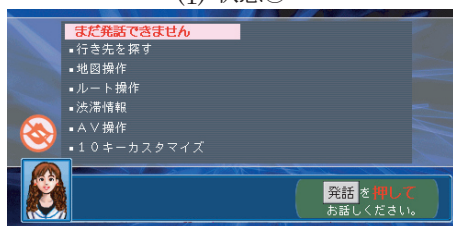


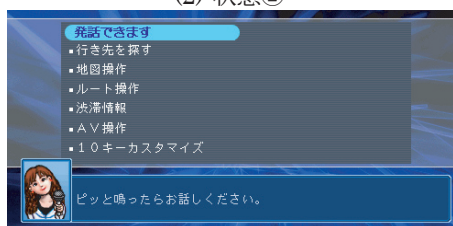
図4 シーケンス改良パターン1のカーナビ音声操作シーケンス



(1) 状態①



(2) 状態②



(3) 状態③

図5 表示改良案4

表示できなくなったコマンドは,"その他"と発話された際に表示するようにした。

また, 表示の改良に関しても, 前記表示改良案3に, 発話できないことを連想する大きなマークをリストの文字に重ねて表示する表示改良案4を追加した。シーケンス改良案1と表示改良案4を併用した場合の表示を図5に示す。状態番号は, 図3のフローに示した番号に対応する。

上記3つのシーケンス改良案を, 6人の被験者に対し, 1人あたり7つのタスクを与えてテストした(1改良案あたりそれぞれ2名ずつ)。画面表示に関しては各被験者に, 対策無し/表示改良案3/表示改良案4の3種類を行ってもらった。

シーケンス改良案の違いによる評価結果を表6に示す。成功率は, 発話可能状態の回数に対する発話タイミングミスを起こさなかった回数の割合である。

シーケンス改良案1は, タスク1では成功率が低いものの, それ以降のタスクでは他の手法より成功率が高くなっている。初めて使う際には何度かミスをする, そこでシーケンスを学習し, 以降の成功率が上がるものと思われる。総発話回数でも, シーケンス改良案1が最も少なく良好な結果となっていることがわかる。よって, シーケンス改良案1が改善方法としては良いと判断した。

次に, 表示改良案の違いによる発話タイミング成功率を集計した。その結果を表7に示す。

表示改良案4が最も良い結果となった。また, 被験者へのアンケートを行ったが, 表示改良案4のマークの表示はわかりやすかったとの意見が多かった。

よってここでは, 発話タイミングミス軽減効果の

表6 シーケンス改良案の違い簡易評価結果

シーケンス改良案	発話タイミング成功率 (%)		総発話回数 (回)
	タスク1	タスク2~7	
案1	57.1	95.4	59
案2	80.0	81.0	89
案3	90.9	84.7	142

表7 画面表示の改善一簡易評価結果2

表示改良案	発話タイミング成功率 (%)
従来方法	81.4
案3	86.2
案4	90.1

最も高い手法候補として、シーケンス改良案 1 と表示改良案 4 との併用を選択した。

3.3 走行時のシーケンス

走行中に音声操作を行う場合、安全性の観点から、現在はリスト表示を行わない仕様になっている。このため、走行中はリストからコマンドを目で探すことはない。その場合、シーケンスとして従来通りシステム主導で発話タイミングを切り替えた方がよいか、停車時と同様シーケンス改良案 1 のユーザ主導で切り替えた方がよいか、簡易実験により確認した。ここでは、4 人の被験者に対し、5 つのタスクを与えた。前半 3 タスクは上記 3.2 で絞り込んだシーケンス改良案 1 と表示改良案 4 との併用である。後半 2 タスクは、2 名は従来シーケンスを、残り 2 名はシーケンス改良案 1 をテストした。双方のシーケンスとも後半 2 タスクは、発話タイミングを通知するマークは表示するがリスト表示はしない。しかし、認識可能なコマンドを読み上げるヘルプ機能があることを、被験者には事前に通知した。また、走行状態の負荷をシミュレートするため、後半 2 タスクでは被験者は PC のドライビングゲーム (自由に走行できる練習モード) を、ステアリングとブレーキ / アクセルで操作しながらタスクを実行した。

結果は、走行時を模した後半 2 タスクでは、双方の方法とも発話回数が増加し、発話タイミングだけでなくタスク遂行が困難な状態となった。これは、後半 2 タスクではリスト表示が無いので、発話する言葉を探すために、ヘルプ機能を多用しているためである。ヘルプ機能では発話可能なコマンドを次々とシステムが読み上げるが、被験者は何度もこのヘルプ機能を実行している。被験者によると、読み上げられたコマンドの中から目的のコマンドを聞きとることが困難なため、とのことであった。

今回我々は、走行中にリスト表示を行えばこのような問題は解決できるのではないかと考えた。表示されるリストは文章ではなく単語で、文字数も少なく視認性が良い。このため、画面表示による運転への影響も少ないと思われる。よってここでは、走行中も停車中と同様、リスト表示を行いシーケンスの変更もしない方法を選択することとした。

4. 効果確認実験

4.1 実験方法

上記のように、発話タイミングミスを軽減するた

めの対策方法として、発話ボタンによるユーザ主導での発話タイミング決定と、画面表示によるシステム状態の通知の併用を選択した。この対策方法の効果を確認するために、12 名の被験者に対し 5 つのタスクを与え、ユーザビリティテストを実施した。5 つのタスクの内後半 2 タスクは、3.3 で行ったように PC のドライビングゲームを操作しながら、タスクを実行した。12 名の被験者のうち、6 名は対策方法を、他の 6 名には従来方法のテストを実施した。テストの後、全ての被験者に対策方法・従来方法の双方を体験してもらい、インタビューを行った。なお対策方法では、2 度目の発話ボタンが押されてから、約 1 秒後にピープ音を発し発話可能状態に切り換えるようにした。

4.2 実験結果

表 8 に、全被験者の発話タイミング成功率の平均を集計した結果を示す。タスク 1 を除き、従来方法に比べ対策方法の発話タイミング成功率が高く、対策方法が有効であることがわかる。対策方法は、タスク 1 での発話タイミングミスが多くなるが、これは 3.2 での実験と同様の傾向となっている。ドライビングゲームを行いながらのタスク 4・5 では、対策方法の発話タイミング成功率が若干劣化するが、従来方法と比べると高い数値を示している。

続いて表 9 に、1 被験者当たり 1 タスクの平均発話回数を示す。

タスク 1 では、従来方法・対策方法双方とも、発話回数が多くなっている。これは、システムの認識できない語彙外発話が多いことによる。タスク 2・3 ではそれが改善されている。ドライビングゲームを行いながらのタスク 4・5 では、従来方法の発話回数が多い。これは、3.3 での実験で確認している通り、ヘルプ機

表 8 対策方法と従来方法の
発話タイミング成功率 (%)

方法	タスク 1	タスク 2・3 平均	タスク 4・5 平均
従来	71.9	67.4	76.6
対策	62.8	91.8	84.1

表 9 対策方法と従来方法の
平均発話回数 (回 / タスク)

方法	タスク 1	タスク 2・3 平均	タスク 4・5 平均
従来	16.0	3.9	23.2
対策	8.5	4.1	3.7

能を多用しているためである。しかし、対策手法ではその傾向が見られず、良好な結果となった。被験者はドライビングゲームの操作を行いながら画面を確認して、音声操作を行うことができていた。

次に、テスト後の被験者へのインタビューの結果と考察を行う。まず、対策手法で発話機会毎にボタンを押すシーケンスに関して、70%の被験者が面倒と答えた。ユーザ主導での発話タイミング決定は、効果があるものの、慣れてくると面倒と感じてしまうようである。発話タイミングを通知する画面表示に関しては、リスト上の文字表示が有効と答えた被験者が70%、マーク表示が50%、キャラクタは0%であった(複数回答を許した)。このことから今回の対策手法は有効であり、文字とマークの2通りの対策を施したことが効果的であった。また、走行中のリスト表示に関しては、90%の被験者が必要と答え、10%の被験者は危険なので必要ないと答えた。今回はPCゲームを用いての模擬走行実験ではあったが、多くの被験者が走行中にリストを表示することは、危険と感じるより必要と感じたようだ。最後に、発話ボタンを押してから発話可能状態に切り替わるまでのディレイに関しては、90%の被験者が長すぎると答えた。

5. まとめ

カーナビ音声操作機能のユーザビリティテストを実施し、認識阻害要因の調査を行った。その結果、発話タイミングの取りづらさが一つの大きな要因であると判明し、対策を検討した。カーナビ音声操作のプロトタイプツールを作成し、対策方法を考案し、プロトタイプを作成し、少人数被験者へのテストを繰り返し行い、対策手法の検討・絞り込みを行った。その結果、画面の改良およびユーザ主導での発話タイミング切り替えを効果のある対策手法として選択した。この対策手法のプロトタイプに対して、一般被験者へのユーザビリティテストを行い、発話タイミングミス軽減効果があることを確認した。

今後、上述の結果に基づいて、製品への導入を検討するとともに、もう一方の大きな課題である、語彙外発話に対す対応の検討も行っていきたい。

参考文献

- (1) 黒須正明: ユーザビリティテストング, 共立出版 (2003)
- (2) 庄境 誠, 中村 哲, 鹿野清宏: フレームワイズな音声検出に基づく適応フィルタを利用した自動車内でのロバスト音声認識, 信学論 (D-II), vol. 81-D-II (No.6), pp.1074-1083 (1998).
- (3) 國枝伸行, 野村和也, 中村一啓: ガイダンス音声の伝搬経路特性変化にロバストなバージョン機能を持つ音声認識システム, 信学技法, SP2003-101, pp.41-46 (2003).

筆者紹介

- 外山 聡一 (そとやま そういち)
技術開発本部モバイルシステム開発センター
- 川添 佳洋 (かわぞえ よしひろ)
技術開発本部モバイルシステム開発センター
- 小林 載 (こばやし はじめ)
技術開発本部モバイルシステム開発センター
- 藤田 育雄 (ふじた いくお)
MBG技術統括部
- 金子 忠靖 (かねこ ただやす)
MBG事業企画部
- 大杉 淳 (おおすぎ じゅん)
MBG事業企画部
- 塩田 岳彦 (しおだ たけひこ)
技術開発本部モバイルシステム開発センター