

E-041

カーナビにおける音声インタフェースの実車評価

In Car Evaluation of Speech Interface on Car Navigation System

相原 政徳 小田嶋 和幸 畑岡 信夫 (東北工業大学)
Masanori Aihara Kazuyuki Odashima Nobuo Hataoka

1. はじめに

早稲田大学との共同研究として、早稲田大学で製作した音声認識ソフトウェアである「FlexibleShortcuts+SelectAndVoice」[1]を用いて音声インタフェースを実車で評価した。評価のタスクは、カーナビでの音声入力機能である。

FlexibleShortcuts+SelectAndVoice の操作は、音声入力とコントローラでの操作入力の両方ができる。また、操作履歴と発話履歴がPC内に記録される。この機能により、語彙外発話や誤認識理由を考察する事が出来た。評価内容は、語彙外発話に焦点をあて、語彙外発話が起きた理由と、語彙外発話が起きた時の被験者の動作、次の行動を解析し事務室での評価結果と比較考察した事前のユーザーへの操作手順(インストラクション)の良し悪しで語彙外発話の頻度と音声認識率に大きく影響することを報告する。

2. 評価実験装置及び評価環境

2.1 場所

東北工業大学駐車場で乗車し、評価実験を行う。

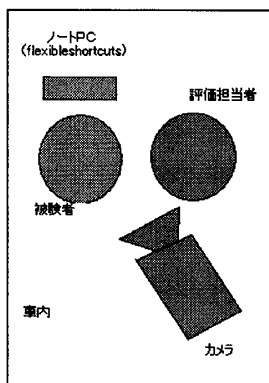


図1 実験場所

2.2 使用器具

使用するアプリケーションは、早稲田大学が構築したPC内のアプリケーションを使用する。図2のコントローラでのコマンド選択と音声認識でのコマンド選択のどちらでも入力出来る。

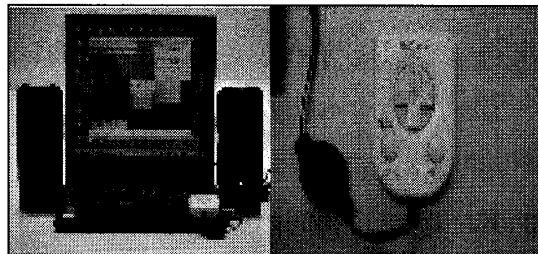


図2 使用器具

2.3 騒音

実車で評価だが、実際に走行するのは危険なためエンジン音や空吹かし等のアイドリング雑音の混じった環境で行う。

2.4 被験者

操作経験のない10名のログを取り、評価する。

2.5 発話条件

各被験者に計10箇所の住所を検索させる。内2箇所は市町村合併などで住所変更されていて、検索出来ない住所となっており、この2箇所の住所検索では語彙外が発生する。住所検索画面では、下記の図の様に、上から都道府県、市区町村、地域、番地と入力する場所が決まっており、違う場所で発話すると語彙外発話となってしまふ。

3. 評価実験方法

3.1 実験手順

- (1) まず初めに被験者に車に乗ってもらい操作方法の説明をする。
- (2) 5分間、評価担当者が、被験者に操作説明をしながら実際に FlexibleShortcuts の操作をする。この際、「コントローラの操作説明」「コントローラの発話ボタンを押しながら話す」という説明をする。また、5分程度、どのような機能があるのかを実際の FlexibleShortcuts の画面上で説明し、計10分間被験者に操作説明をする。

被験者に練習のため、自由に操作してもらふ。評価担当者がアドバイスをしながら、被験者にアプリケーションの操作をさせ、例題として東北工大八木山キャンパスの住所(宮城県仙台市太白区八木山香澄町1-1)を検索させる。更に語彙外の例として(宮城県古川市荒谷1-1→宮城県大崎市古川荒谷1-1)を検索させる。これにより、被験者に市町村合併等による語彙外があることを認識してもらふ。

また、このような語彙外が発生したとき本当に語彙外なのかを確認するために住所検索のタスクから一覧を見ることが出来ることをしっかり教える。これにより、被験者が実際の実験で語彙外が発生した場合の対処法を身につけさせる。

被験者が、操作を出来るようになったら本実験に移る。

(3) 本番

本番ではアドバイス無しで行う。実験中のPC画面を評価のためにビデオで録画する。実際に車に乗っているという想定なので空吹かしは2000~4000回転で行う。実験の始まりと終わりの合図はビデオの担当者が指示を出す。実験内容は、被験者に計10箇所の住所を検索させる。内2箇所は市町村合併などで住所変更されていて、検索出来ない住所となっており、この2箇所の住所検索では語彙外が発生する。実験の目的として誤認識理由、語彙外が発生したときの被験者の挙動を調べる。

4. 評価結果

4.1 発話の比率

表1に、実車と事務室での評価結果を示す。[2][3]

表1 評価結果

実車評価	事務室評価	
	回数	割合
認識	505	64%
語彙外	149	20%
誤認識	126	16%
バグ	0	0%
計	780	953

4.2 語彙外発話回数と種類

語彙外発話の種類は、住所検索タスクでは、1.語彙外住所、2.住所の区切り誤り(政令指定都市)、3.入力場所、4.トップ画面 FlexibleShortcuts タスクでの語彙外発話の4種類があった。表2語彙外発話の回数と種類をまとめた表で、上が事務室での評価結果、下が実車での評価結果である。

	語彙外住所	住所の区切り	入力場所	トップ画面	計
実車	84回	42回	17回	6回	149回
事務室	112回	55回	21回	49回	237回
	48%	23%	9%	20%	

表2 語彙外発話

4.3 語彙外発話後の被験者の動作

語彙外発話の種類ごとに、語彙外発話後の被験者の動作を評価した。

語彙外住所

被験者に10箇所の住所を検索してもらい、その中の2箇所に市町村合併などで住所変更されていて、検索出来ない住所を検索した時の被験者の行動を評価する。検索出来ない住所だと気付くまでに発話した回数を図4に示す。

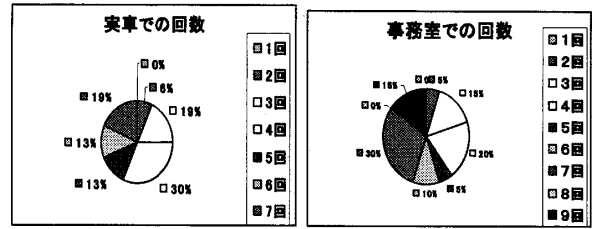


図4 語彙外住所までに気付く回数

5. 考察

実験前の結果予想では実車による測定の方が事務室での測定結果より認識率が下がると考えていた。しかし、結果的には実車での測定結果の方が語彙外発話や誤認識が少なかった。認識率は実車で64%事務室で52%、語彙外が20%と25%、誤認識が16%と23%であった。このように全ての面で実車での測定の方が上回る結果となった。

このような結果になった理由として、被験者の違い、評価担当者の説明の違いによるものが影響していると考えられる。被験者の違いでは個人の能力差でも結果が大幅に変わってくると思われる。勘がいい人だと語彙外住所だと3~4回目まで気付く場合が多いが、認識されるまで何度も発話する人もいた。

次に操作手順の説明(インストラクション)の良し悪しに影響される。事務室での評価の時の説明は比較的簡単に説明していたのに対して、実車での評価では10分間丁寧に説明し実験した。(実車での評価実験のときは語彙外住所があるかもしれないということを事前に被験者に伝えておいたが、事務室での評価実験の際はしなかった)実際に被験者に対し音声認識について聞いてみると「認識しない」「使い勝手が悪い」など良いイメージをもたれてないのがほとんどだった。しかし、今回実験をしてみて、「しっかりと説明があったことで意外に使いやすくイメージが変わった」と使い勝手が悪いという最初のイメージから変わったのが分かる。つまり、認識率を上げたり語彙外発話や誤認識を少しでも減らす方法の大きなポイントはしっかりとしたインストラクションだと考えられる。

6. まとめ

今後の課題として、インストラクションの質を上げて音声認識の使い勝手を良くすることと、語彙外発話に対しては特にシステムからの注意を与えるなどの配慮が必要である。

参考文献

- [1]早稲田大学 小林研究室「音声インタフェース構築支援技術及びランダムモニタリング技術に関する研究開発」、経済産業省「音声認識基盤技術の開発」平成19年度成果報告書 PP-140-171(平成20年3月)
- [2]東北工業大学 瀧直人 笹原大亮 卒業論文「カーナビにおける音声インタフェースの評価」(平成21年3月)
- [3]瀧直人, 笹原大亮, 畑岡信夫, 中野鐵兵, 熊井 朋之, 小林哲則: カーナビにおける音声インタフェースの評価— 語彙外発話の状況と対応案に関して —, 信学技報, Vol. 108, No. 465, SP2008-152 (2009年3月)