

多言語会話文、語彙データを利用した 災害救援者教育用アプリケーションの開発

平松 初珠[†] 石島 悌[†] 萬宮 健策[‡] 山根 聡[‡] 堀 一成[‡]

大阪府立産業技術総合研究所[†] 大阪大学[‡]

1. はじめに

近年、世界で未曾有の大災害が起こっている¹⁻²⁾。2009 年だけでもインド南部の大洪水、スマトラ沖地震、サモア沖地震などの災害が発生している。このような災害が発生した際に、日本から自衛隊やボランティアスタッフなどが現地へ赴き、救援活動を行っている。その際、被災者や現地の救援スタッフとの言葉の壁の解消が、大きな課題となっている。

大阪大学では、専攻 25 言語の基礎的資料収集とそのデータベース化、さらに教育への応用を大きな目標として掲げている。ペルシア語専攻やウルドゥー語専攻で、2005 年のパキスタン地震の際に救済者のための語彙集を作成し、ウェブサイト³⁾に公開した実績がある。このウェブサイトは、災害時の救援活動参加者から高い評価を得た。

これを基に、大阪大学では、現地語に不慣れた日本人が救援活動を行う際に必要な会話文や語彙データを選定し、整備する取り組みを継続して行っている⁴⁻⁶⁾。また、ウルドゥー語だけではなく、アラビア語、タイ語でも同様の取り組みを行っている。蓄積しているデータは、現地語やカタカナ表記の発音といった、言語データである。また、文字が読めなくてもコミュニケーションが図れるように、言語データだけではなく、仕草や音声、画像データも蓄積している。文字情報だけでは得られないこれらの情報を蓄積していることは、文化学習という点からも有用な情報であり、このことは大きな特徴である。

今回、被災地で手軽に利用できる日本語に対する現地語などを表示するアプリケーションの開発について紹介する。このアプリケーションに、上述の蓄積したデータを搭載する。また、アプリケーションを実際に使用した学生の感想を報告し、最後に今後の展望を述べる。

2. アプリケーションの開発

アプリケーションを搭載する携帯端末は、被災地での利用を想定し、小型で堅牢なコンピュータを選定した。OS は、Windows Vista である。なお、開発は、Microsoft Visual Basic 2008 Express Edition で行った。

また、本アプリケーションは、クライアント・サーバモデルではなく、コンピュータ上にすべてのデータを搭載し、一台で完結して動作する構成にした。これは、被災地などインターネットの利用が困難な場所での使用を想定したためである。

本アプリケーションの構成を図 1 に示す。基本構成は、実行ファイルと設定ファイル、および実行に必要な 3 つのライブラリファイルである。これらの容量は、あわせて 1 メガバイト以下である。これにデータを加えたものが全体の容量となる。これは、データの容量にもよるが、全てを USB メモリや SD カードに十分収めることが可能な容量である。

設定ファイルには、1 行に 1 言語の情報を載せている (図 2)。順に、後に示す言語データの csv ファイル名及び画像と映像ファイルの入ったフォルダ名、アプリケーション上で表示される言語名、現地語の文字の表示方法を表している。3 つ目の項目は、文字を右から書く言語があるため、書字方向を数字で区別する。

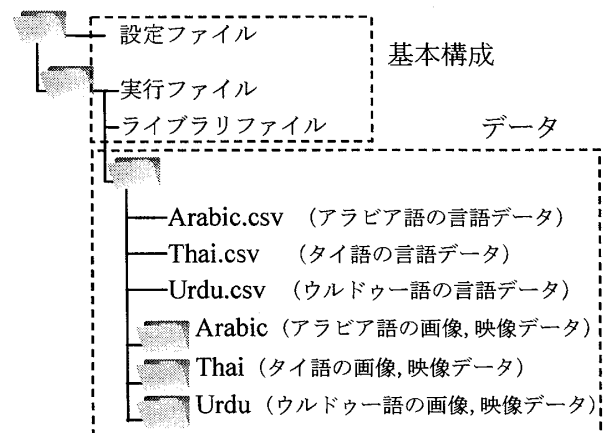


図 1. アプリケーションの構成

Development of Educational Application for Disaster Relief
using Multilingual Conversational Sentences and Lexical Data
[†]Hatsumi HIRAMATSU, Dai ISHIJIMA
Technology Research Institute of Osaka Prefecture
[‡]Kensaku MAMIYA, Sou YAMANE, Kazunari HORI
Osaka University

#通し番号, 日本語, 現地語, カタカナ表記の発音, 画像ファイル名, 映像ファイル名

1, ありがとう, شکریہ, シュクリヤ, 3020297.JPG,

2, チクッとします, 我慢してください, ذرا درد ہو, صبر کیجیے, زلزلہ ڈالڈ ہور, سابل کیجیے, tyuusya.bmp, 14.wmv

図 3. ウルドゥー語言語データ (Urdu.csv) の一部

[LANGUAGE]		
Arabic, アラビア語, 1,	#右横書き (←)	
Thai, タイ語, 0,	#左横書き (→)	
Urdu, ウルドゥー語, 1,	#右横書き (←)	

図 2. 設定ファイル

データは、設定ファイルの情報を基に、図 2 に示すような構成で、言語単位で管理している。言語データは、言語ごとに csv 形式で保存している。音声や画像の非言語データは、言語ごとに作成したフォルダに語彙単位で保存している。次に、csv ファイルの例を図 3 に示す。図 3 に示す順番で、1 行に 1 語彙の情報を載せている。これらの項目は、全てを用意する必要はない。例えば、図 3 の 1 番目は、映像ファイル名がない。対応するデータがない場合は、その項目が表示されないだけである。

言語ごとにデータの管理を行っているので、言語の追加は比較的容易である。最低限必要な作業は、設定ファイルに言語情報を 1 行追加し、データを用意するだけで、新たな言語が追加される。

開発したアプリケーションは、誰でも簡単に操作できるような設計にした。アプリケーションの画面例を図 4 に示す。使い方は、言語を選択し日本語のデータ一覧から利用したい語彙を選択するだけである。語彙を選択すると、その日本語に相当する、現地語、カタカナ表記の発音、画像、映像（音声）の少なくとも 1 つが表示される。

本アプリケーションは、日本語からそれに対応する現地語を検索するだけでなく、現地語からそれに対応する日本語を検索できる機能を有している。この双方向の検索機能により、日本人と被災者や現地の救援スタッフなどとの双方向コミュニケーションをより円滑に図ることができる。その他、選択する語彙を絞り込む機能や、新しい語彙を登録する機能などを実装している。

3. 使用した感想

実際にウルドゥー語のデータを搭載したアプリケーションを学生に利用してもらい、感想を聞いた。映像を利用できるのは興味深い、使い方によってはゲーム感覚で使うこともでき親しみがわく、といった肯定的な意見があった。また、会話文の内容が限定的なので、日常生活で用いられる表現なども整備して搭載してみたいという提案があった。

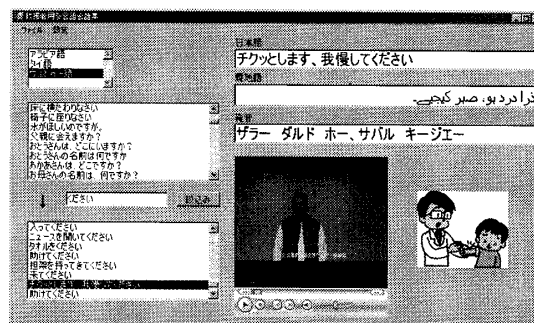


図 4. 画面例

4. まとめと今後の展望

大阪大学が整備したデータを、被災地で手軽に利用できるアプリケーションを開発した。また、開発したアプリケーションを使用した感想を聞いたところ、おおむね肯定的であった。今後は、対応する語彙の増加を図るとともに、より汎用性のあるアプリケーションへの改良を行いたい。

また、本アプリケーションは、現地語から日本語を検索する機能を有しているため、日本人が被災地での救援活動に役立つだけでなく、逆に、日本で災害があった際に、海外からの救援者とのコミュニケーションを補助するツールにもなる。さらに、災害に限らず、空港、薬局、観光地、日常生活など幅広い分野での応用が期待される。今後、これらの分野の語彙、会話文の充実、また、対応する言語の増加も図る予定である。

謝辞

本研究は、平成 19～21 年度 文部科学省 科学研究費 萌芽研究 課題番号 19652058（代表者：堀一成）の補助を受けています。

参考文献

- 1) 内閣府（編）：平成 21 年度防災白書，佐伯印刷（2009）
- 2) 「Asian Disaster Reduction Center (ADRC)」，<http://www.adrc.asia/top_j.php>，2010/01/08 アクセス
- 3) 「大阪大学外国語学部ウルドゥー語専攻 | 救援者のための会話」，<<http://www.sfs.osaka-u.ac.jp/user/urdu/homepage/zalza/trans.html>>，2010/01/15 アクセス
- 4) 山根 聡：「災害救援者支援のための会話集等作成について」，大阪大学世界言語研究センター論集 1 号，(2009)，pp.217-225.
- 5) 萬宮健策：「インターネットを利用した新たなウルドゥー語教育」，インド・パーキスタン語学科開設 100 周年記念シンポジウム，（東京外国語大学，2008 年 12 月）
- 6) 堀一成，山崎直樹，竹原新，小島一秀：「多言語並行マルチメディア言語資源の構築」，言語処理学会第 13 回年次大会発表論文集（2007），pp. 768-771.