

# インターネットを利用した音声のききとり公開実験の試み(1)

## 5F-10 - システムの構築に関する技術的検討と稼働状況 - \*

○足立隆弘 山田玲子 駒木亮 片桐滋  
ATR 人間情報通信研究所

### 1 はじめに

近年、個人でのインターネット接続数が急速に増加していると共に、WWWの利用が一般的になりつつある。しかし、インターネットの利用はこれまで情報の提供者から消費者に対しての一方的なものがほとんどであり、ネットワークの双方向性を実用的な面で不特定多数向けに活用した事例はあまり多くない。本研究では、音声知覚用実験アプリケーションの一部をこれらのインフラを利用して公開実験という形で提供することにより、従来研究室や実験室の中だけで行われてきた実験を新しい形で行うことを試みた (<http://bluebacks.hip.atr.co.jp> [3]にてデモ)。具体的には、我々の研究所で、言語音獲得過程を明らかにするために研究されてきた日本語話者による米語 /ɛ/ /ɪ/ 音のききとりという興味深い問題を公開実験の課題としてとりあげた。本報告では、この公開実験の概要と稼働状況を報告する。また、このような公開実験では技術的な面で様々な問題が発生する可能性がある。これまで発生した問題およびそれらに対する対策に関しても言及する。

### 2 問題となる事柄およびその対策

Webベースでの公開実験を行う際に、問題となる事柄は次のようなものが挙げられる。

#### 2.1 実験用ソフトの問題

聴覚実験の方法としては、オンラインで実験する方法と、必要なソフトやデータ等をローカルに置いて結果のみを送るオフラインで実験を行う2つの方法が考えられる。そこで今回の公開実験に先立って、前者をCGI,Javaを用いて実際の実験の環境として耐えうるかを検証した。その結果、通信速度の悪化に伴って実験が成り立たなくなる可能性があることが分かった。また、対象となる一般のモデムアクセスでの参加者にとっては、参加費用(電話代、ISP利用料等)の増大が考えられた。したがって、本研究ではオフラインでの実験を選択することにした。

一方、ローカルで実験を実施する場合には、実験用ソフトの配布方法の問題が発生する。本研究では、一般向け書籍付録[1]としての配布および、Web上からダウンロードの二通りの方法で配布した。

#### 2.2 インターフェイスの問題

配布した実験用ソフトは我々が実際に聴覚研究に用いているソフトをベースに開発したものである。通常実験室内で実験を行う際には、試験者が被験者に対してインストラクションを行う。また、結果のログは試験者が解析に適した形で書き出

される。しかし、今回のような公開実験の場合は、参加者の計算機操作に関する習熟度にばらつきが予想されたため、一連の操作を容易に行えるようにした。また、ログファイルは圧縮した形でサイズの小さいテキストファイルに書き出すように変更した。

実験結果を登録する場面でも、Webの平易なインターフェイスを用い、簡単な操作で登録や結果の閲覧の操作が出来るようにした。また、登録時にはいくつかの個人的な情報が併せて登録される。この部分では個人情報の取り扱いには留意することを明示することにより、登録に際して不必要な心理的障壁を設けないように心がけた。

登録した結果については、登録者個人の結果と全登録者の平均が、グラフを多用した分かりやすい形式で表示/反映されるような構成にした。また、その表示のタイミングも、登録直後に即座に表示されるような構造にした。

これらのことを行った結果、参加者が実験を実施、結果の登録、閲覧を迷わずに操作することが可能なインターフェイスを実現出来た。

#### 2.3 サーバー側処理のパフォーマンスの問題

公開実験を行うに際し、実際にどのくらいのアクセスがあるかは全く未知数であった。そのため、身近にあるハードウェア構成で、出来るだけ高いパフォーマンスを発揮可能なものにし、参加者に対して不満を感じさせないようなレベルにするために、以下のことを考慮した。

##### 2.3.1 ハードウェアおよびOSの選択

Web serverとして使用する環境として検討対象となったマシンは、中程度の性能を持つSPARCstationおよびIBM PC/AT互換機(以下PC)である。このようなサービスを提供するサーバーとしては、WorkStation(以下WS)を使用するのが一般的である。しかし、PCおよびWSのパフォーマンスをByteBenchmarkで調べた結果、図1のようになった。構成によってこの値は変動するが、PCがWSに対してパフォーマンスで遜色無いことが分かる。特に、今回のような重い数値演算を行わないような用途に対しては、十分実用的である。そこで、本公開実験ではPCを使用することにした。

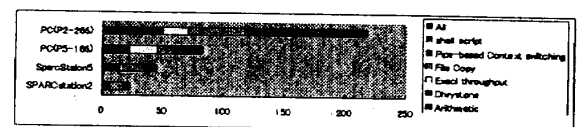


図1: Byte benchmarkの結果

ハードウェアとしてPCを選択したことから、OSはOpen-SourceのLinuxやFreeBSD、SolarisやSCOのような商用UNIX、そしてWindows95やWindowsNTのようなWindowsベースのOSが選択可能となる。まず、Windows系はメ

\*Internet inquiry on perception of non-native speech contrast  
(1) System development: by Takahiro Adachi, Reiko A-Yamada, Ryo Komaki and Shigeru Katagiri(ATR Human Information Processing Research Laboratories, Seika-cho, Kyoto 619-0288, Japan), <http://www.hip.atr.co.jp/>

メンテナンスに手間がかかり、また、パフォーマンスもそれ程発揮できないために選択肢から外した。Open SourceのUNIXと商用UNIXとの比較であるが、研究所内での実績や多くのWeb上の報告などを考慮した上で、Linuxを選択した。Linuxは軽いOSとして実績がある。同一のプラットフォーム上でSolarisを動かした場合に比べ、はるかにパフォーマンスが良いことが報告されていることが選択した理由の一つである。

このようなことを検討した結果、Web serverはPCを用い、OSとしてLinuxを使用することにした。使用したPCの諸元は、Pentium 166MHz, RAM 96MByteである。

### 2.3.2 Webサーバーの構成 (Software)

UNIX上では商用ソフトとOpen Sourceのソフトとそれ程大きなパフォーマンスの差がないことが多い。そこで全面的にOpen Sourceのソフトを使用し、その上にHTMLドキュメントおよび各種CGIプログラムを構築することにした。

データの登録や参照等はSQL DBを用いることが開発効率上好ましい。今回のようなケースでは同時に複数のアクセスが予想されるため、thredに対応したMySQLを選択した。また、MySQLは使用ライセンス上の制約が少ないことも、メリットの一つとしてあげられる。

CGIプログラムは基本的にはFORMのハンドリングはperlで記述し、DBのアクセスや各種統計計算等のパフォーマンスを要求される部分はCで記述した。なお、ダイナミックにグラフを生成する部分にはgdライブラリを使用し、DBのデータを元に、GIF画像のグラフを生成するようにした。

### 2.3.3 ページ構成の問題

多くのサイトで、デザイン性を重視した画像データを多用する重いページが増加してきている。しかし、今回の実験ではモデムでアクセスするようなユーザーが多いことが予想されるため、重いページは好ましくない。一方、多くの参加者を得るためにはデザイン性や閲覧性、操作性は重視すべきである。そのため、作成するドキュメントに関しては、テキストを中心にした軽いページとすると共に、デザイン的にもアトラクティブになるように心がけた。

## 3 公開実験の途中経過

### 3.1 アクセス数の推移、データの登録数

公開実験は実験用ソフトを付録として添付した書籍の刊行日と同じ、1998年3月20日に開始した。4月には新聞などのマスメディアで紹介されたこともあり急速にアクセスが増加し、6/30までの3ヶ月間で累計199,657ヒットを記録した(図2)。

7月現在は毎日500ヒット前後で推移している。なお、1日辺りの最高ヒット数は4/2の30,291ヒットであり、データ登録数は現在1220人を越えている。jpドメインのみのアクセス数とデータの登録数の内訳を、図3に示す。本実験の参加者の半分以上は、企業、学校からではなく、プロバイダを経由してアクセスをしている個人であることが示されている。

### 3.2 サーバーの稼働状況

3/20の立ち上げ以来の4ヶ月間、マシンのリプレース及び電源点検のための停電以外全くトラブル無く連続稼働し、ソフト的/ハード的なトラブルは皆無であった。また、当初懸念していた各種の処理も全く問題ない速度で動作しており、データの閲覧やグラフの自動作成等もリアルタイムで表示されている。

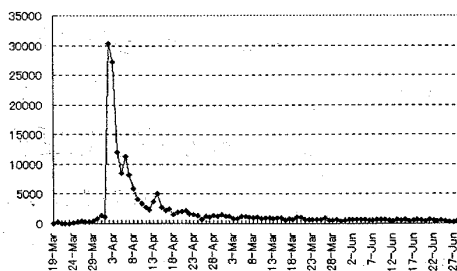


図2: アクセス数の推移

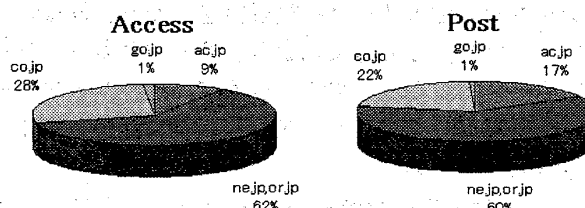


図3: アクセス、登録の内訳

## 4 考察

今回の公開実験では多くの参加者を得ることが出来、このような形態での実験が十分現実的なことが明らかになった。また、プラットフォームとして全てをOpen Sourceの環境で行い、PC上で稼働させたとしても、十分実用的なものになりうることを示した。さらに、ベンチマークテスト等の結果、性能の面では商用のソフトやWSの方が劣る場合もあった。ソフトの開発効率が高いことも考慮すると、本システムで用いた環境は、コストパフォーマンスの面では特筆すべきものがある。

また、登録されたデータを解析した結果、日本語話者にとってどのような/e//i/音の特徴がききとり困難かが明らかになった。同時に、これまで被験者として実験室での実験に参加しもらうことが困難であった層も個人として多く参加しており、実験対象を広げることが出来た。

本実験で得られたデータの信頼性については、実験室での統制された条件下でのデータとの比較検討を行った。その結果、上記の点に着眼する限りは信頼できる結果が得られていることが明らかになった(本予稿集 駒木他 5S-02)。

今後、多言語間での比較を行うため、本Webページを多言語化すると共に、本試みから得られた技術的ノウハウを生かしたインタラクティブな情報提供サービスや、異なる実験を行う予定である。

## 参考文献

- [1] 山田恒夫・足立隆弘・ATR人間情報通信研究所: 英語リスニング科学的上達法, 講談社, 98-03-20
- [2] 駒木亮 山田玲子 足立隆弘 片桐滋: 5S-02 インターネットを利用した音声のききとり公開実験の試み(2). 情報処理学会 98秋全国大会, 98-10
- [3] 足立隆弘 山田玲子 駒木亮 片桐滋: デモ6 インターネットを利用した音声のききとり公開実験の試みデモ. 情報処理学会 98秋全国大会, 98-10