

## 主観評価実験による音声入力とキー入力との比較

三宮, 真智子  
鳴門教育大学

高木, 英行  
松下電器産業株式会社中央研究所

樺澤, 哲  
松下電器産業株式会社中央研究所

三尾, 忠男  
鳴門教育大学

<https://hdl.handle.net/2324/4479719>

---

出版情報 : 日本音響学会講演論文集, pp.109-110, 1987-10. The Acoustical Society of Japan  
バージョン :  
権利関係 :

# 3—5—10 主観評価実験による音声入力とキー入力との比較

△三宮真智子<sup>1)</sup> ●高木英行<sup>2)</sup> 樺澤哲<sup>2)</sup> △三尾忠男<sup>1)</sup>

<sup>1)</sup>鳴門教大

<sup>2)</sup>松下電器 中研

## 1. はじめに

現実の音声認識技術は理想からはまだ程遠いが、実用化の観点からすれば、入力者がキーボードより好ましいと判断する技術レベルに達する事が、当面の目標である。筆者等は音声WPを例に取り、主観評価実験を通じてこの技術目標を探ってきた。その結果、単音節入力がキー入力に勝るまでの技術ギャップが相当大きいこと<sup>1)</sup>、文節音声入力ならば近い将来この技術目標に到達でき得る可能性があること<sup>1)2)</sup>、が明らかになった。本報ではこれらの結果を基に、文節音声入力のWPがキー入力のWPより優位になる音声認識率を主観評価実験により求める。

## 2. 評価実験

**2.1 概要** 実験は、キー入力と文節音声入力(平均音節認識率: 65%, 70%, 75% のいずれか)とで55文節の文章を入力し、評定尺度評価を行なうという方法を用いた。実験の状況を図1に示す。WPは音声認識シミュレータ(パソコン上に実現)に接続され制御される。実験者2はキー操作・発声を監視指導する。実験条件を表1に示す。

表1. 実験条件

比較対象:	キー入力 × 文節音声入力
入力方式:	かな入力 × 文節連続音声
音節認識率:	単音節単位で 65%, 70%, 75%
入力文章:	心理学の児童書から抽出した55文節・239音節
被験者:	各音節認識率毎に8名(男2女6) 計24名
評価方法:	9項目に対する5段階の評定及び内観報告の分析

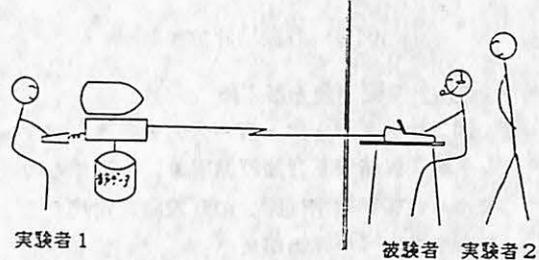


図1 実験状況

**2.2 入力方法** 液晶7行表示のWPを用いる。キー入力は「かな入力」で行ない、音声入力は文節単位に連続発声する。音声入力時には文節音声認識シミュレータを用い、認識率を制御しながらWPに表示する。誤り傾向は実際の単音節認識装置に合わせてある。また本シミュレータは言語処理部を含み、発声者には文法的に正しい候補のみ提示する。誤認識の場合は、訂正キーで順次候補を表示し正しい文節を得る。候補にない場合は、再発声するか諦めてかなキー入力する。なお、本報での音声認識率とは言語処理前の値であり、第1候補認識率を音節換算したものである。

**2.3 被験者** 3種類の音声認識率各々に8名(男2女6)の学部学生計24名を被験者とした。被験者はキーボード経験が浅く、かな入力には慣れていない。

**2.4 評価方法** 図3の評価9項目に対して図2の5段階評定を行なう。実験手順は、

- (1) キーで文章入力、
  - (2) 音声で文章入力、
  - (3) 両入力に対する評定、
  - (4) 実験を通しての内観報告、
- である。

\* "Subjective Assessment of Voice Entry and Keyboard Entry," by M.Sannomiya<sup>+</sup>, H.Takagi<sup>++</sup>, S.Kabasawa<sup>++</sup>, T.Mio<sup>+</sup> (+Naruto Univ. of Education, ++Matsushita Elec. Ind. Co., LTD.)

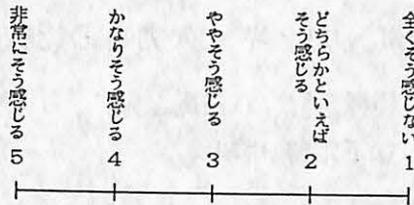


図2 5段階評定尺度

### 3. 実験結果

図3は、キー入力・音声入力の各々に対する主観評価結果を音節認識率毎に示したものである。横軸数字は図2の5段階評定尺度であり小さい方が評価が良い。

### 4. 考察

t検定を用い、キー入力と音声入力との有意差検定を各評定項目毎に行なった。結果を図3の右端に示す。◎は有意差あり、△は有意差傾向あり、と言える。本評価実験結果から音節認識率が75%になれば、音声入力の方がキー入力より好ましいと評価されると判断できそうである。

入力条件の一定なはずのキー入力において部分的な評価結果のバラつきが見られるが、実験時期・時間帯によって被験者間に疲労感等の差が若干生じたためではないと思われる。この影響はキー入力・音声入力の両方に及んでいるはずであり、両者間の評定差は有効な情報として信頼できると考えられる。

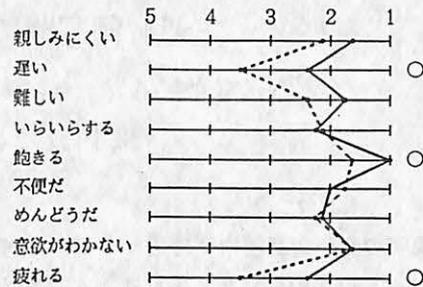
### 5. まとめ

音声認識の実用化の観点から、音声WPを例に取り音声入力達成すべき認識率を明らかにした。その結果、文節単位に連続発声し、音節認識率75%を達成すればキー入力より高く評価されることが明らかになった。

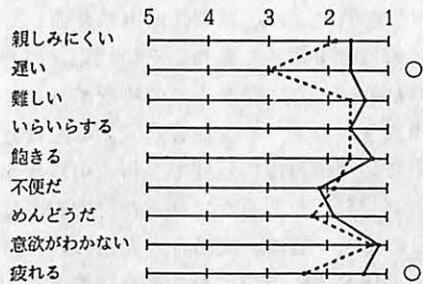
今後、実験の評価9項目が総合的な好ましさを十分表現しているかどうかを因子分析等を通して検討し、実験結果の裏付けをする予定である。最後に関係者各位に感謝致します。

### 参考文献

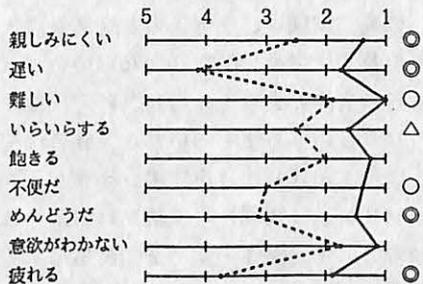
- 1) M.Sannomiya, et al. "A Study on Voice Recognition as Human Interface for Japanese Word-Processor," 第2回ヒューマン・インタフェース・シンポジウム 論文集 pp545-552, Oct.1986
- 2) 高木, "音声入力ワードプロセッサの主観評価," 第2回ヒューマン・インタフェース・シンポジウム 論文集 pp357-360, Oct.1986



(a) キー入力 × 音声入力 (音節認識率65%)



(b) キー入力 × 音声入力 (音節認識率70%)



(c) キー入力 × 音声入力 (音節認識率75%)

図3 キー入力(.....)と音声入力(——)との主観評価結果

- ・上部の数字は図2の5段階評定尺度
- ・右端の記号はt検定による有意差
  - ◎: 1%の危険率で有意差あり
  - : 5%の危険率で有意差あり
  - △: 10%の危険率で有意差傾向あり