

音声認識ビジネスの現状と将来展望*

西村雅史

日本アイ・ビー・エム 東京基礎研究所

E-mail: nisimura@jp.ibm.com

1. はじめに

音声認識技術を利用した音声ビジネスへの期待が高まってから相当時間がたつが、少なくとも日本ではまだ大きなビジネスには育っていないようである。98年ころディクテーションソフトが各社から販売されると、ある種のブームともなり、音声認識がPCビジネスの中で大きなウェイトを占めた時期もあったが、それも一時的なものであった。はたして音声認識技術は、世の中に大きな革新をもたらすことが出来るのか、それとも結局使い物にならない技術のまま終わってしまうのだろうか。

音声認識ビジネスの現状について概観するとともに、今後期待される応用分野として、車載機器のインターフェース、コールセンタービジネスへの活用について、現状技術の問題点と、今後の見通しについて述べる。

2. 音声認識ビジネスの現状

「安全、効率、便利」といった観点で、既存のインターフェースを置き換える形での音声インターフェースが、次のような対象に対し古くから検討されて来た。

- 1) パソコン(キーボード、マウスの代替)
- 2) 携帯情報端末(10キー、十字キー、ペン等の代替)
- 3) 電話系サービス(プッシュホンの代替)
- 4) 車載情報機器(リモコン、タッチパネル等の代替)

いずれの場合も、既存のインターフェースに対する優位性が常に問われることになる。

パソコンの音声認識ソフトや、携帯電話の音声オートダイヤルなどは、一時的にはビジネスに貢献したが、標準インターフェースにはならなかった。カーナビ用の音声認識機能も、自動車メーカーの純正品を中心にして搭載例は多いものの、お世辞にも広く使われているとは言い難い状況にある。一方、電話系のサービスにおいて、プッシュホンによる自動応答に代わる手段としての音声認識は、ある程度の成功を収めている。日本では、証券会社の株価照会システムが、その成功例として知られている。この例では、1) 複雑な条件設定が不必要、2) 株価を知ることだけが目的であり、誤認識に対して生じるコストが小さい、という好条件が成功につながったと考えられる。ただ、テレフォンショッピング、チケット購入、契約内容変更など、この他のコールセンター業務の自動化では苦戦も目立つ。

一方で、汎用インターフェースとしてよりも、音声認識による音声のテキスト化機能に着目した応用例も多い。特に、電子カルテの普及にとともに、医療所見や看護記録の入力など、医療用の応用が根強い支持を得ている。ただ、医療機関ごとに使う語彙や表現が異なるなど、普及を妨げる要因もまだ多い。また、録音機器の操作性や、マイクの装着性といった副次的な理由で敬遠されることもある。

この他にも、

- 1) テレビ番組やビデオ教材等の字幕作成支援
- 2) 議会などでの議事録作成支援
- 3) 音声データ検索
- 4) 外国語教育

などの応用例が知られている。特に、ビデオ教材などに対する字幕作成は、技術的制約はあるものの、すでに広く使われている。ただ、いずれもそれほど大きなビジネスにはまだ育っていない。

* Current and future trends in speech recognition market, Masafumi Nishimura, Tokyo Research Lab., IBM Japan Ltd.

3. 音声認識ビジネスの可能性

音声認識技術がビジネス上、大きな価値を生むためには、

- 1) 顧客の満足度向上
- 2) ビジネスの効率改善(コスト削減、売り上げ増加)

のいずれかに貢献することが重要である。

ユーザーインターフェース技術としての効率改善は、安全性の向上、入力効率向上などを通じて顧客(あるいはユーザー)の満足度向上に直接つながるはずである。逆に言うと、顧客を満足させることが出来ないレベルの技術では意味がない。現状の音声認識技術は、認識エンジンの性能ばかりが注目され、システムの使い勝手という点では作りこまれていない場合が多い。アメリカではいくつかの電話音声自動応答システムが顧客の高い評価を受けていると聞くが、認識エンジン自体の性能というよりも、認識誤りへの細かい対応など、システムとしての完成度が評価されている。

また、音声認識技術を単なるインターフェース技術としてではなく、ビジネスの効率改善のための手段という立場で捉えることも重要になってきた。たとえば、音声認識技術とベクトル空間モデル等の情報検索技術にもとづくコールルーティング(問い合わせの電話を担当部門に割り振る業務)の自動化は、コールセンターにおけるコスト削減、効率向上に大いに寄与している。

以下では今後期待される車載情報機器インターフェースとしての可能性、ならびに、コールセンタービジネス拡大の鍵となる可能性について紹介する。

3.1 車載情報機器インターフェースとしての可能性

現在、車の中で最も音声認識が利用されているのは明らかにカーナビの操作に対してである。ハンズフリー、アイズフリーを実現するインターフェースとして、注目される一方、リモコンやタッチパネルで操作可能な項目の多くが音声操作可能であったとしても、音声認識を利用するユーザーは少ないと聞く。ある程度有効利用されているのは目的地設定の際のPOI(Point of Interest)入

力ぐらいである。現状の車載音声認識装置はスイッチの置き換えに過ぎないが、今後対話型のインターフェースが導入され、自動化のレベルが上がった場合に、ユーザーは無条件で音声認識を使うようになるだろうか。

我々は、都市圏在住のドライバー100名に対し、将来の音声対話インターフェースに関するアンケート調査を行った。具体的な状況と、対話例を示してのアンケートである。これによると、窓やエアコンといった機器の操作、あるいは観光案内といったものは、たとえ高度な音声対話が実現されても利用したくないという結果が多数を占めた。スイッチの代替とはなりえないことは当然としても、観光案内などについても、走行中に無理に音声を使って不確かな操作を行うより、あらかじめ出発前に調べておけばよいことだと考える人が多いようである。このように、音声対話インターフェース(あるいはテレマティクス)に対する期待度は総じてあまり高くない。

一方で、音声に期待する項目としては、緊急性の高い項目、主に非常時の対応と渋滞情報の入手が高い割合を占めた。特に、適切な渋滞情報の入手とその活用については要望が高く、この点では今後のITSの高度化ともあいまって、音声対話インターフェースがビジネスにつながる可能性は十分にある。

3.2 コールセンタービジネスへの応用

近年、コールセンターの機能は飛躍的に向上しており、CTI(Computer Telephony Integration)、IVR(Interactive Voice Response)、音声録音装置などが多くの企業に導入された。それと同時に、コールセンターは従来の“お客様相談室”のレベルから、CRM(Customer Relationship Management)戦略の拠点へと役割を変えつつある。この変化にともない、音声認識技術に求められる役割も、単なる自動応答技術から、オペレータのコンプライアンス管理、音声対話録音の書き起こし、そして、それに基づくコールセンターログの作成や、マーケティング活動に利用するためのテキストマイニングなどに広がると予想される。

まず、自動応答に関しては、すでにIVRの導入はコールセンターの57%に達し、音声認識装置の導入も12.6%に達したという報告がある¹⁾。この報

告によると、24時間対応の要求は多く、どの業種でもIVRの利用率は高い（IVR利用者の内訳は93%がプッシュホン、音声認識は7%）が、そのうちIVRの再度の利用に抵抗感を示さなかった人は43%のみである。残りの57%はIVRの再度の利用に抵抗感を示した。理由は、「人より時間がかかる」「操作が面倒」「質問できない」などであり、IVRは必ずしも顧客満足度向上にはつなげない。

操作性の向上が急務だが、プッシュホンを置き換えただけの音声認識ではその役割を果たしておらず、高度な音声対話技術に期待がかかる。IBMでは音声認識と自然言語理解技術との組み合わせにより、高度な対話システムを構築し、高い顧客満足度を得ているが、一方で、システム構築に膨大な費用と時間がかかるという問題があった。ビジネスとして成功させるためには横展開を容易に実現できるようなTOOL群を整備できるかどうかは鍵となる。

なお、サポートセンターでのIVRの利用も多いが、これはコールルーティングへの利用が多い。ただ、この場合も過半数がプッシュホンによるIVRへの対応を面倒と感じている。先に述べたように、これについては今後音声認識技術の利用が期待できる。

コールセンターでは音声録音装置の導入も進んでおり、すでに54%のコールセンターで採用されている¹⁾。将来、高い精度で対話音声の自動書き起こしが実現できることが条件となるが、この録音をテキスト化し、オペレータの発言のモニタリングに活用することが考えられる。顧客が、今後のコールセンターに期待する項目の上位には、電話の待ち時間短縮、24時間対応に加え、オペレータの教育が入っており¹⁾、オペレータの適切な教育は顧客満足度の向上にもつなげる重要なポイントである。また、コンプライアンス・チェックの観点からもこの応用に対する期待は大きい。さらに顧客側の音声も含めて自動書き起こしが可能になれば、テキストマイニング技術との融合により、コールログに代わるマーケティング活動の重要な情報源になると期待される。なお、この応用において、キーワードによる音声データ検索技術はすでに実用可能なレベルにあるとも言えるが、この方法では管理者はまだ膨大な量の音声

を聞きなおす必要があるので効率は必ずしも良くない。

一方、ビジネス的にはインバウンドだけでなく、アウトバウンドコール（電話セールス）に対する自動化の期待も高いと聞くが、そもそもアウトバウンドコールに対しては顧客の96%が不快感を示すという結果もでている¹⁾。インバウンドコールからの適切な情報抽出と、マーケティング戦略が必要である。

4. おわりに

音声認識技術に対する期待はまだ総じて高いものの、現状の延長線上にビジネスチャンスがあるかということ必ずしもそうではない。激化する世界的な技術競争に勝ち抜くためにも、本当に社会の役に立つ技術とはなにかということをもう一度真剣に考えてみるのもいいだろう。

謝辞

車載情報端末における音声対話インターフェースの有効性に関する調査結果を提供いただいた伊東伸泰氏に感謝いたします。

[参考文献]

1) “コールセンター白書2004”、コンピューターテレフォニー編集部編、リックテレコム（2004年7月）