

第5回 SLP デベロッパーズフォーラム

イベント企画の趣旨について

～コーディネータの立場から～

庄境誠¹

音声認識、音声合成、音声対話を始めとする音声技術は、ハイテク技術が市場で普及する上で一般的に避けられないと言われていた落とし穴、いわゆる「キャズム」を越えたであろうか、あるいは、近い将来、越えることができるだろうか？ 今回の第5回 SLP デベロッパーズフォーラムでは、音声技術が「キャズム」を越えて普及の段階に入り、産業化を実現する上での課題とその解決方法に関して、事業面及び技術面から議論したい。

Message from a Coordinator of events in the 5th SLP Developer's Forum

Makoto Shozakai¹

Have speech technologies, covering speech recognition, speech synthesis and speech dialog, passed over the trap, so-called "Chasm", which is generally said inevitable for the high-tech. to become popular in the market? Or, will those be possible to pass over it in near future? It is advised that we discuss about problems and their solutions to make the speech technologies widely spread and to create the industry from business and technological points of view during events in the 5th SLP Developer's Forum.

1. はじめに

音声認識技術、音声合成技術、そして、音声対話技術などの音声技術への過去の研究開発投資により、音声技術のここ10年の進歩は目覚ましい。しかし、その技術進歩が、音声事業、音声産業の創出に結実していないという点に関しては、同意される方も多いのではなからうか。音声技術は、人工知能技術の一環として、過去も現在も脚光を浴びているが、夢やビジョンが先行し、収益性のある事業には未だ到達していないのが実情である。この現実には、音声技術に立脚した製品やサービスの事業化を目指す企業にとっては、厳しい状況である。その理由、解決課題、解決策についての議論は過去何度も繰り返されて来ているが、残念ながら全てが解決されたわけではない。世の中の状況や音声技術を取り巻く事業環境も時々刻々変化しており、その折々において、今後も継続的な議論が必要である。

この業界に身を置く者は皆、音声技術の普及と産業化を願っていると思うが、それを果たすためには、単に技術的な側面ばかりでなく、事業的な側面に関しても十分に検討する必要がある。マーケティングの世界では、ハイテクが普及する上でどうしても避けられない落とし穴として、いわゆる「キャズム」の存在が知られている。今回の第5回 SLP デベロッパーズフォーラムでは、音声認識を始めとする音声技術が、この「キャズム」を越えるための課題とその解決策について、事業面及び技術面から議論したい。

2. 音声技術のキャズム

たまたま、ふとしたことから、ジェフリー・ムーア著「キャズム～ハイテクをブレイクさせる「超」マーケティング理論」¹⁾という本を読む機会があった。この本は、ハイテクのマーケティング理論のバイブルとも評され、今なお、多くの読者を魅了し続けている。

Chasm (キャズム) とは、ハイテクが陥り易い、不可避の落とし穴のことである。以下、この本の記述・用語に従って、キャズムについて概説する。

一般に、ハイテクのユーザは、以下の5つのカテゴリーに分類される。

- 1) Innovators (ハイテクオタク)
- 2) Early Adopters (「ビジョン先行」派)
- 3) Early Majorities (「価格と品質重視」派)
- 4) Late Majorities (「みんな使ってるから」派)
- 5) Laggards (ハイテク嫌い)

キャズムとこれら5つのカテゴリーの関係を図1に示す。

[†] 旭化成株式会社 新事業本部 情報技術研究所
Information Technology Laboratory, New Business Development, Asahi Kasei Corporation

ハイテク製品は、先ず、ハイテクに飛びつく Innovators に魅力的な機能仕様を備えなければならない。次に、ハイテクにより、他社（者）に仕様で差をつけることを目的とする Early Adopters に採用されなければならない。Innovators や Early Adopters は数の上では少数である。これらのニッチ市場を最初に顧客とすることができなければ、そもそも、そのハイテク製品の将来性はないと言っても過言ではない。しかし、これらのニッチ市場では、Innovators や Early Adopters の高い期待に応えることが優先されるために、多機能となり、当然のことながら高価格にならざるを得ない。Innovators や Early Adopters は、高価格でもハイテク製品を購入するが、十分な開発・評価期間を費やす時間的余裕がないために、製品としての完成度が低く、品質は不十分であることが多い。そのため、市場からは製品に対する不満が多く寄せられることになる。また、高価格であるために、製品の販売台数は伸び悩む。従って、Innovators や Early Adopters を対象とする間は、製品の研究開発投資を回収することが難しい。この状態が長く続くと、そのハイテク製品はその製品寿命を終えることになる。

一方、Innovators や Early Adopters 向けのニッチ市場において、製品開発をする中でハイテク技術は、ある部分技術において、成熟の域に達する。その成熟した部分技術を活用すれば、より限定した機能で、低価格、高品質の新たな製品を開発販売することができる。そのようなハイテク製品は、価格と品質を重視する Early Majorities に訴求する。さらには、Early Majorities が使っているから自分も使う Late Majorities の購買意欲も刺激することができる。Early Majorities と Late Majorities の主ターゲットを顧客として得ることにより、ハイテクはようやく研究開発投資の回収サイクルを回すことができるようになるのである。

ところで、ユーザのある少数は、ハイテクが嫌いな Laggards と呼ばれる。これらのユーザはハイテクの対象顧客ではないので、考慮の対象外である。

Early Adopters と Early Majorities の間に待ち構える不可避な落とし穴、裂け目を Chasm（ギャズム）と呼ぶ。ギャズムを越えるには、製品開発上の発想の転換と人材の交代が必要とされる。ハイテクにとって、このギャズムを越えられるかどうかは死活問題と言われている。実は、先に紹介した本の帯には以下のように記されている。

“ギャズム”を越えられなかった（越えていない）ハイテク技術&サービス&製品：
電気自動車、ニューラルネットワーク・ソフトウェア、ビデオ会議、オブジェクト指向データベース、人工知能（AI）、DAT、音声認識、インターネット端末、衛星携帯電話、OS/2、ISDN、キャプテンシステム、ペンコンピュータ、Newton、デスクトップ・パブリッシング・・・・・・・・

この本が著されたのは 90 年代であり、今から 10 年以上前のことである。しかし、音声技術ベンダーの一人として事業を冷静に分析した場合に、音声技術は果たして「ギャズム」を越えることができたのであろうか、あるいは、できるのであろうかという不安がどうしてもよぎる。

表 1 に、車載機器における音声認識機能仕様の変化についてまとめた。従来は、他社にカタログ仕様で勝つことが優先された時代であった。Innovators や Early Adopters を対象とした高価格のハイエンドカーナビが製品化されてきた。(1)コマンド制御、(2)電話番号や電話帳の登録名でのダイアリング、(3)住所や施設名（POI(Point of Interest)）での目的地登録、(4)アーティスト名、アルバム名、曲名、ジャンル名による音楽検索などの多くの機能が盛り込まれてきた。ただし、正直なところ、音声認識技術の実力から背伸びした多機能、高機能仕様であったため実力が伴わず、結果として、実環境での性能が仕様未達成のケースがあった。また、多くの計算量やメモリ量を必要とする高価格な製品にならざるを得なかったため、販売台数も伸びなかった。

昨年 9 月のリーマンショックに端を発した、100 年に一度と言われる経済危機の影響は未だ続いている。音声認識技術の最も典型的な商業的応用例の一つが、カーナビなどの車載機器のハンズフリー音声インターフェースであるが、全世界的な自動車販売の不振により、車載機器の販売台数も落ち込んでいる。この事実は、当然、音声ベンダーの経営にとっても脅威である。

しかし、今回の経済危機は、従来の社会のシステムやものの考え方を大きく転換させる機会と言われている。それは、車載機器の音声インターフェースの製品ラインナップの考え方にも影響を与えている。具体的には、ローエンドカーナビ、ディスプレイオーディオ、Bluetooth ハンズフリーなどの、Early Majorities と Late Majorities のための安価な車載機器製品の開発を促している。そして、それらに搭載される音声認識仕

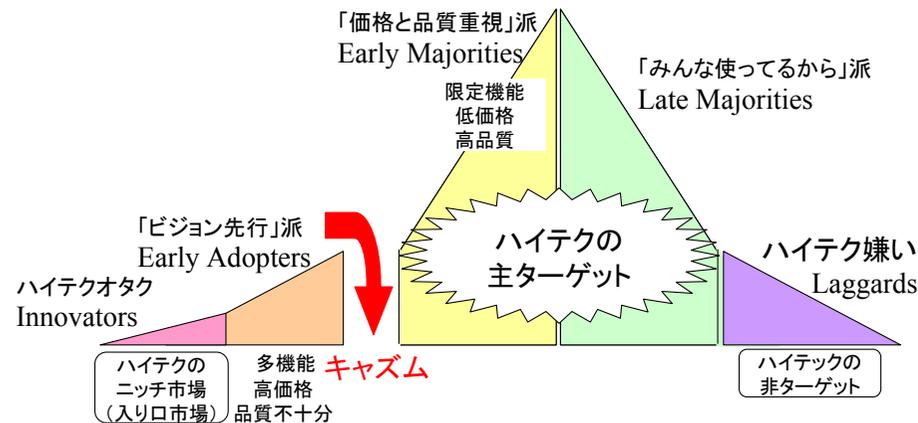


図 1 ハイテクの落とし穴「ギャズム」

表1 車載機器における音声認識機能仕様の変化

		従来	今後		
		ハイエンドカーナビ	ローエンドカーナビ	ディスプレイオーディオ	Bluetoothハンズフリー
↓音声認識機能	→価格	高価格	中価格	中価格	低価格
(1)コマンド制御					
(2)ダイヤリング (電話番号/電話帳)					
(3)目的地登録 (住所/施設名)					
(4)音楽検索 (アーティスト名, アルバム名, 曲名, ジャンル名)				(iPOD)	

様は、既に開発されたハイエンドカーナビに搭載された音声認識機能のサブセットになっている。この動きにより、まさに、上記の「キャズム」理論の通り、安く買えて、真に役に立つ音声認識の需要と普及が本格化することが期待できる。つまり、音声認識が「キャズム」を越える好機が到来したとの見方もできるのである。

さらに、以下の好情報も有る。北米では、販売される車種の80%に標準またはオプションとして音声認識技術による音声入力機能を有する車載機器製品が用意されており、その製品の購入率は40%強と言われる。一方、日本でも、販売される車種の40%に同様の製品が用意されており、その製品の購入率は50%強と言われている。これまでは、音声認識技術による音声入力機能を有するカーナビは上位車種にのみ設定されることが多かった。従って、高価なカーナビにしか音声認識機能が搭載されていない現実があった。しかしながら、時間の経過とともに、音声認識機能、音声インタフェース機能の完成度が高まったのは事実である。今後は、中位から下位の車種にも音声認識機能、音声インタフェース機能が展開される流れである。これもまた、音声技術が「キャズム」を越える見込みがある根拠の一つと言える。

こうして、音声技術は、間違いなく、機能競争から性能品質競争に移行する。

3. イベント企画の構成

今年で第5回目を迎えるSLPデベロッパーズフォーラムのイベント企画のコーディネータを務めるに当たっては、音声技術が「キャズム」を越えるための方策のヒント

が得られるような議論の場をアレンジしたいと考えた。

そこで、日本における音声認識ベンチャーの草分け的存在である、(株)アドバンス・メディア社の藤田氏に、「音声認識 実用化事例の紹介とその課題」と題して、招待講演をお願いした。電子カルテの音声入力、議会の議事録作成、コールセンター通話の文字化などの具体的な実用化事例に関して、御紹介いただけるものと思う。

また、2006-2008年度にかけて行われた、経済産業省研究開発プロジェクト「情報家電センサー・ヒューマンインタフェースデバイス活用技術・音声認識基盤技術」(以下、国プロ)の参加機関の中から、早稲田大学の中野氏、三菱電機(株)の石川氏、旭化成(株)の岡本氏、(株)日立製作所の大淵氏に御発表をお願いした。

まず、中野氏には、「音声インタフェース普及促進のため開発支援技術」と題して、効果的な音声技術の使い方や評価技術という観点を中心に、音声認識アプリケーション開発を支援する技術について語っていただく予定である。

次に、石川氏に、「音声UI設計とその評価～普及促進にむけて、システム構築の立場から～」と題して、音声UIのユーザビリティの向上、開発コストの低減、ユーザ中心の評価とそれに基づく改良、標準化・共通化の必要性に関するこれまでの種々の議論を踏まえ、国プロの成果、車載機での内外の動向、今後のサービス、その実現手段に着目しつつ、開発者視点で音声インタフェースの普及促進について考察し、課題と可能性を示していただく。

また、岡本氏には、「人間の多様な振る舞いを考慮した音声UIの必要性」と題し、国プロの成果を踏まえて、どのような点に注意したら「使える音声UI」をデザインできるかについて紹介してもらうとともに、認識率を上げるだけではなく、リアルタイムで音声UIに対するユーザの多様な振る舞いをモニタリングし、計測したユーザ習熟度に合わせて音声UIの動作が変化していく技術の必要性を指摘してもらう。

さらに、大淵氏には、「実世界に存在する音声・音響を対象とした認識技術」と題して、実世界に既に存在する多くの音データを用いた分類・検索技術のメリットについて紹介いただくとともに、音声と非音声の判別、ログデータからの音声検索、音響イベントのセンシングといった応用の観点から、音声・音響処理の新しい方向性について語っていただく予定である。

そして、これらの御発表の後に、「音声認識技術がキャズムを越えるには？」と題したパネルディスカッションを計画している。上記の藤田氏、石川氏、大淵氏に加え、日本IBM(株)の西村氏にもパネリストとして御参加いただいて、事業面、技術面など様々な角度から忌憚のない議論をしていただく予定である。

参考文献

- 1) Geoffrey A. Moore, "Crossing the Chasm," 1991, 1999, (日本語版)キャズム～ハイテクをブレイクさせる「超」マーケティング理論～, 川又政治訳, (株)翔泳社