

## 音声 HMI 技術・製品の品質の計測，評価，保証

庄境誠<sup>1</sup>

音声 HMI 技術・製品の普及を日本で実感する機会はまだ多くないが、海外の、特に自動車環境では、安全や Mind Distraction, Telematics Service の観点からハンズフリー音声 HMI は既に実用化・普及の段階に達し、直感的で使い易い対話型音声 HMI としての高い完成度や品質が要求されるようになった。日本においても、自動車環境だけではなく、モバイル環境、住宅環境などでのユニバーサルインタフェースとして、音声 HMI 技術・製品の普及が期待される。そのための、音声 HMI 技術・製品の品質の計測、評価、保証技術の確立の必要性、課題と解決方法などについて論じる。最後に、今回の第 6 回 SLP デベロッパーズフォーラムのイベント企画の内容を簡単に紹介する。

### Measurement, Evaluation and Assurance of Quality of Speech HMI Technologies and Products

Makoto Shozakai<sup>1</sup>

There are less opportunities where we can feel spread of speech HMI technologies and products in real life. On the other hand, hands-free speech HMI has already reached to phase of realization and spread in terms of safety, mind distraction and Telematics services especially in automotive environment abroad. High completeness and quality of an intuitive and easy-to-use interactive speech HMI is now highly demanded. In Japan, the spread of speech HMI as universal interface in not only automotive environment but also mobile environment and home environment is expected. This paper discusses about necessity of technology establishment of measurement, evaluation and assurance of quality of speech HMI technologies and products. Their problems and solutions are suggested. Finally, the arranged presentations in this 6th SLP Developer's Forum are briefly introduced.

## 1. はじめに

音声認識技術、音声合成技術、雑音除去技術、音声対話技術などの音声 HMI のここ 10 年の技術的な進歩は目覚ましい。音声 HMI 技術・製品の普及を日本で日常的に実感する機会はまだ多くないが、海外の、特に自動車環境では、運転行動中の安全や Mind Distraction の回避、移動中の Telematics サービス提供の観点からハンズフリー音声 HMI は、既に実用化・普及の段階に達し、直感的で使い易い対話型音声 HMI としての完成度や品質が要求されるようになった。例えば、既に、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 HMI を有するカーナビやカーオーディオなどの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% に同様の製品が提供されており、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 HMI 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 HMI 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 HMI が展開される流れである。また、今年の 4 月には、車の運転中、携帯電話や音楽再生機器などの操作を音声 HMI を通じて行っている運転手は、手で操作する運転手に比べて道路から目を離すわき見が極端に少なく(10-40%)、従って、Mind Distraction を大幅に回避できるため、近い将来すべての自動車メーカーが音声 HMI を車に搭載するようになるというバージニア工科大学交通研究所 (VTTI) の調査結果が発表された<sup>1)</sup>。今後は、自動車環境だけではなく、モバイル環境、住宅環境などでの、若年層から高齢者層までのユニバーサルインタフェースとして、音声 HMI 技術の普及が期待されている。

昨年の第 5 回 SLP デベロッパーズフォーラムでは、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、マーケティングの世界でよく知られている「キャズム(Chasm)<sup>2)</sup>」について紹介し、音声 HMI 技術も他のハイテクと同様にこの「キャズム」を免れることが出来ず、それを乗り越えてさらなる実用化・普及の局面に到達す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<sup>3)</sup>。

そこで、本稿はその議論を継承して、音声 HMI 技術・製品の品質の計測、評価、保証が「キャズム」を乗り越えるためのキープポイントであることを指摘し、その技術確立の必要性、課題と解決方法などについて論じる。

最後に、今回の第 6 回 SLP デベロッパーズフォーラムのイベント企画の内容を簡単に紹介する。

<sup>†</sup> 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

## 2. 音声 HMI 技術・製品の現状認識

一般に、ハイテクのユーザは、Innovators(ハイテクオタク)、Early Adopters(「ビジョン先行」派)、Early Majorities(「価格と品質重視」派)、Late Majorities(「みんな使ってるから」派)、Laggards(ハイテク嫌い)の5つのカテゴリーに分類される。Early Adopters と Early Majorities の間に待ち構える、ハイテクが陥り易い不可避の落とし穴(裂け目)はキャズムと呼ばれる。(キャズムの説明を最後に詳述するので参考にされたい。) Innovators, Early Adopters に歓迎される多機能、高価格、品質不十分の小市場の段階から、このキャズムを越えて、Early Majorities, 続いて Late Majorities に受け入れられる限定機能、低価格、高品質のボリュームゾーンの段階に移行できるかがハイテクにとって死活問題と言われており、製品開発上の発想の大きな転換が必要とされている。

音声ベンダーの1社として、音声 HMI 技術・製品市場の現状を見ていると、他のハイテクと同様に、音声 HMI 技術・製品もこのキャズムを越える苦しみを味わっているように見える。音声 HMI 製品の機能や価格に関する発想の転換が必要であることは言うまでもないが、本稿では特に、キャズムを越えるポイントとして、音声 HMI 技術・製品の品質の計測、評価、保証に焦点を絞って論じたい。

数年前、ある顧客のマネジャーから言われた一言が今でも強く印象に残っていて耳を離れない。その言葉は、「音声認識のことを知れば知るほど、使い方が難しい」というものであった。これはまさに、音声 HMI 技術・製品の品質の現状を言い表している。音声 HMI 技術・製品の品質保証技術が確立されていないことに対するベンダーへのクレームであると理解した。つまり、顧客は概して、音声認識、音声合成、雑音除去、音声対話などの音声 HMI に係る技術の

A)個別品質

B)統合品質

を、ベンダーが計測し、評価し、顧客に対して保証することを求めている。しかし、従来は、その点が不十分であり、結果的に音声 HMI 技術・製品の品質の計測、評価、保証に要する顧客側でのコストが高むことに対する改善要求がベンダーに突きつけられている。

音声 HMI 技術・製品の品質の計測、評価、保証に関する技術に関する研究開発の努力は、これまでももされてきたが、確立されているとはとても言い難く、業界標準規格として明確化されているわけでもない。この点は、今後、音声 HMI 技術・製品が普及していくためには、学会・業界として乗り越えていかなければならない課題である。複雑な音声 HMI 技術・製品であればあるほど、その品質の計測、評価、保証に要する技術は複雑化し、技術の確立に必要な研究開発の努力の道程は長く、多大な工数やコストの負担を余儀なくされる。先ずは、シンプルな仕様の音声 HMI 技術・製品を対象

として、上記 A)個別品質や B)統合品質の尺度を定めることから始めてはどうだろうか。

例えば、音声認識を例にした場合、A)個別品質の尺度として、従来は、平均認識率のみが議論されることが多かった。しかしながら、音声認識機能を搭載した機器を商品化するメーカーは、ユニバーサルインタフェースとしての位置付けを重視するようになってきた。彼らは、機器を販売したユーザ全てに対して高い顧客満足度を提供することを目標としている。従って、話者間や単語間の音声認識性能の平均値のみならず、音声認識性能の分布(ばらつき)にも注目するようになってきている。そこで、例えば、音声認識性能の分布(ばらつき)を表現する評価尺度として、以下を提案する。

Highest(音声認識性能の最高値)

Mean(音声認識性能の平均値)

Mean-1\*sigma(音声認識性能の平均値-音声認識性能の標準偏差値)

Mean-3\*sigma(音声認識性能の平均値-3×音声認識性能の標準偏差値)

10% worst(低認識率話者10%を除外した場合の音声認識性能の最低値)

Lowest(音声認識性能の最低値)

一方、B)統合品質に関しては、個別品質の評価尺度ではなく、ユーザが体感するユーザビリティ(使い勝手)、顧客満足度などの、音声 HMI に対する人間の体感尺度が必要であると思われる。そこで、従来のタスク達成率、タスク達成時間、顧客満足度主観評価値(5段階、7段階)の他に、例えば、習熟度を提案したい。

従来、初心者や高齢者を含む「誰でも」が「簡単に」情報家電機器や車載情報機器を活用できるユニバーサルなインタフェースとして、音声 HMI 技術・製品に安易に期待する向きが多かった。しかし、現在の音声 HMI 技術・製品の能力は未だ人間の音声対話能力には及ばず、ユーザに提供できる機能・性能には制約がある。ユーザはその制約に習熟し、その制約の範囲の中で音声 HMI を使いこなすことが望まれる。しかし、人間が音声 HMI の制約に習熟する速度や程度は様々である。例えば、我々が行ったある被験者評価実験の例では、60代のタスク達成時間の平均値は、20代のその約3.3倍であった。極端にタスク達成時間が長かった68歳男性を除いた場合でも、60代のタスク達成時間の平均値は、20代のその約1.5倍であった。年齢層が高くなるにつれて、タスク達成時間は長くなる傾向があるといえる。また、同じ20代でも、タスク達成時間の最大値は最小値の約2.3倍であった。このようにユーザによって、音声 HMI に対する習熟の速度・程度には大きな差があり、ユーザの習熟度をリアルタイムで計測する技術の実現が待たれる。

この技術が実現できれば、時々刻々変化するユーザの状況に依存した「今だけ、こ

こだけ、あなただけ」のパーソナルサービスを、ユーザが意識して選択するのではなく、ユーザを包み込むアンビエントな音声 HMI 側から提供することが可能となり、ユーザにとって快適な生活を実現できるのではないだろうか。

昨今、HMI の分野では、製品やサービスを利用する過程の品質を重視し、ユーザが真にやりたいことを本人が意識しなくても「楽しく」「面白く」「心地よく」行える点を重視したユーザエクスペリエンスの概念が広がっている。ユーザに意識させることなく習熟度を計測し、アンビエントに音声 HMI をパーソナライズすることができれば、まさにユーザエクスペリエンスの概念に合致しており、音声 HMI 技術・製品は広く世の中に普及するのではなかろうか。

### 3. イベント企画の紹介

今回の第 6 回 SLP デベロッパーズフォーラムでは、音声認識のみに閉じることなく、音声 HMI 技術・製品の全般の品質の計測、評価、保証に関する議論を喚起したいと考えた。その主旨に沿って、イベント企画の内容を企画した。最後に、イベント企画の内容を簡単に紹介する。

まず、西田氏（ATR-Trek 社）に、「分散型音声認識、音声翻訳サービスの現状と展望」と題して招待講演を御願ひした。会社紹介、事業・サービスの概要、携帯電話向け音声認識、音声翻訳サービスの概要と状況、自動音声翻訳実証実験概要、学習による効果、音声認識、音声翻訳サービスの課題と取り組みについて語っていただけるとお聞きしている。

私の本稿（発表）を間にはさんで、音声 HMI の 4 つの分野（雑音抑圧音声、音声合成、残響下音声認識、音声認識サービス・製品）で、その品質の計測、評価、保証に関する発表をコーディネートさせていただいた。

トップバッターは、山田氏（筑波大学）である。「雑音抑圧された音声の主観・客観品質評価法」と題して、雑音抑圧された音声のユーザ体感品質を評価するための、主観・客観品質評価法について語っていただく。

次は、「音声合成の利用シーンと要求される品質との関係」と題して、三井氏（日本電気）に発表を御願ひした。顧客によって、様々な目的や利用環境が存在し、それに応じて、音声合成に求められる品質も異なる。実ビジネスでの経験を踏まえ、求められる品質や、実際に直面した具体的な課題を紹介いただく。

さらに、西浦氏（立命館大学）からは、「室内音響指標に基づく残響下音声認識性能の計測、評価、保証」と題して、残響下音声認識性能を左右する残響環境、スピーカとマイク間の距離、発話者の個人性などの様々な視点で、室内音響指標を用いた残響下音声認識性能の計測、評価、保証を論じていただく。

最後に、今回の第 6 回 SLP デベロッパーズフォーラムのイベント企画が契機となって、今後の学会・業界横断的な議論が活発化し、音声 HMI 技術・製品の計測、評価、保証が実現される日が少しでも早く到来することを期待して止まない。

### 参考文献

- 1) [http://www.usfl.com/Daily/News/10/04/0414\\_028.asp](http://www.usfl.com/Daily/News/10/04/0414_028.asp)
- 2) 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
- 3) 庄境誠, “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～,” SLP78-7, 2009.

(参考) キャズムとは？

キャズム(Chasm)とは、ハイテクが陥り易い、不可避の落とし穴のことであり、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)で最初に提起された。一般に、ハイテクのユーザは、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)の5つのカテゴリーに分類される。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。Early Adopters と Early Majorities の間に待ち構える不可避な落とし穴、裂け目をキャズムと呼ぶ。ハイテクにとって、このキャズムを越えられるかどうか死活問題と言われている。

ハイテク製品は、まず、それに飛びつく 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が、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。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る。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。

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 (キャズム) と呼ぶ。キ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、このキャズムを越えられるかどうか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。

**“キャズム”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、キャブテンシステム、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1社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「キャズム」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。

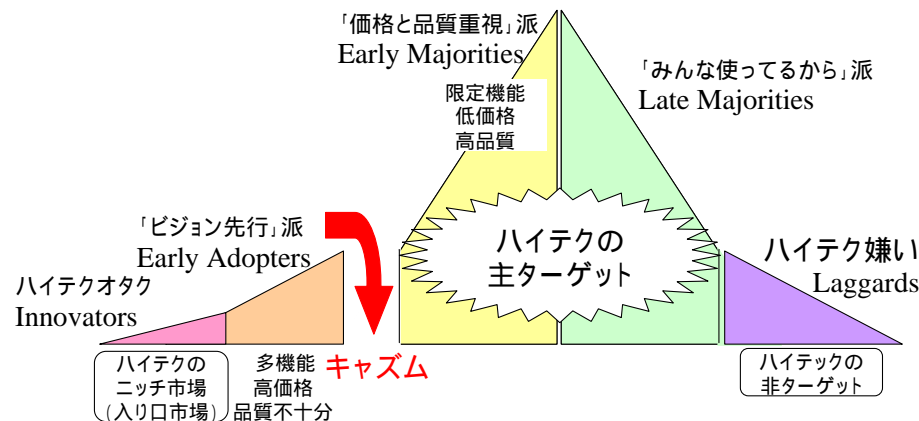


図1 ハイテクの落とし穴「キャズム」