

JEITA 一般社団法人 電子情報技術産業協会

平成 25(2013)年度 音声認識・合成技術 応用製品の動向について

— 音声入出力方式 標準化専門委員会 実施の製品動向調査結果より —

2014 年 11 月

音声入出力方式 標準化専門委員会

○概要

音声認識・合成技術の適用分野の拡大，応用製品の普及を目的として，本委員会では，継続して製品動向調査を実施している。本調査は，応用製品の開発者，製品の導入・購入を考える企業・個人に有益な情報を提供するとともに，継続実施することにより，技術の進歩，利用領域の変遷を理解する貴重な資料となっている。

本年度（平成 25 年度）の調査は平成 22 年度以降の調査方法を基本的に踏襲し，以下の方法により実施した。

(1) 情報提供の依頼

平成 26（2014）年 1 月に，JEITA 本委員会ホームページに，情報提供の依頼を掲示するとともに，音声研究・開発者向けメイリングリスト（onsei-mail@aist.jpn.org）に同一内容のメールを配信。なお，前年に情報を提供いただいた企業には，直接メールを発信し，依頼した。

(2) 掲載製品

情報提供企業が製造あるいは販売する音声認識・合成関連製品であり，平成 25（2013）年 12 月末現在で，日本国内で購入可能な製品とした。

(3) 提供内容

委員会で作成した，製品分類ごとの製品特徴・仕様を記載するフォームを配布，記載いただき，電子ファイルとして受領した。

(4) 掲載内容

提供いただいた情報は基本的に一切の加工を行わず掲載している。ただし，内容の変更を伴わない書式の統一，情報提供者の明らかな誤解による分類誤りについては，委員会にて変更をしている。

記載されている内容については，委員会にて確認をしておらず，内容は，すべて情報提供企業にその責任がある。

○音声合成のシステム開発者向け製品の動向調査結果

日本国内で販売されている音声合成製品を調査した結果について以下に概要をまとめる。
当製品調査は、調査対象の製品分類を大きく

「ミドルウェア組み込み向け」,

「ミドルウェア PC」,

「ハードウェア」,

「その他」

の4カテゴリ（第1分類）に分けた上で、さらに細分類（第2分類）を設けて製品情報をまとめている。

調査結果を表A-1～9に示す。

以下では、その調査結果をもとに、分類ごとの製品件数とその推移について述べる。

表1に、調査対象とした12カテゴリの第2分類のうち、該当製品が存在した9カテゴリについて製品件数と推移を示す。

表1 分類毎の音声合成製品の件数と推移

分類		表番号*	H25 (2013)	H24 (2012)	H23 (2011)	H22 (2010)
第1分類	第2分類		件数(増減**)	件数(増減**)	件数	件数
ミドルウェア 組み込み向け	規則音声合成	表 A-1	38 (+2)	36 (+14)	22	24
	その他	表 A-2	1 (±0)	1 (±0)	1	0
ミドルウェア PC	規則音声合成	表 A-3	29 (+3)	26 (±0)	26	23
	その他	表 A-4	1 (-1)	2 (+2)	0	0
ハードウェア	規則音声合成 (規則音声合成 LSI)	表 A-5	2 (±0)	2 (±0)	2	1
	その他	表 A-6	1 (±0)	1 (±0)	1	2
その他	サーバ製品	表 A-7	10 (-3)	13 (-5)	18	8
	PCソフトウェア ・ツール等	表 A-8	24 (-1)	25 (+2)	23	18
	サービス	表 A-9	9 (-1)	10 (+3)	7	3
合計			115 (-1)	116 (+16)	100	79

*表番号：合成製品一覧の対応表番号、 **増減：前年度からの増減

今年度（平成25年度）の回答製品数は総計115件となり、平成24年度調査と比較して総数で1件の減少となった。

以下、平成24年度と比較した今年度の件数の推移をまとめる。

第2分類でみると、製品件数が増加したのは、

- 「ミドルウェア組み込み向け」の「規則音声合成」製品（2件増）、
- 「ミドルウェア PC 向け」の「規則音声合成」製品（3件増）

であり、逆に製品件数が減少したのは、

- 「ミドルウェア PC 向け」の「その他」製品（1件減）
- 「サーバ製品」（3件減）
- 「PC ソフトウェア・ツール等」（1件減）、
- 「サービス」（1件減）

の4項目で製品数が減少した。それ以外の項目では製品数は同じであった。

表2では、該当製品が存在した9カテゴリについて、製品件数ではなく、製品を提供しているベンダ数とその推移を示す。

表2 分類毎の音声合成製品の提供ベンダ数とその推移

分類		表番号*	H25 (2013)	H24 (2012)	H23 (2011)
第1分類	第2分類		社数 (増減**)	社数 (増減**)	社数
ミドルウェア 組み込み向け	規則音声合成	表 A-1	11 (+1)	10 (-2)	12
	その他	表 A-2	1 (±0)	1 (±0)	1
ミドルウェア PC	規則音声合成	表 A-3	13 (+1)	12 (-2)	14
	その他	表 A-4	1 (-1)	2 (+2)	0
ハードウェア	規則音声合成 (規則音声合成 LSI)	表 A-5	2 (±0)	2 (±0)	2
	その他	表 A-6	1 (±0)	1 (±0)	1
その他	サーバ製品	表 A-7	8 (-1)	9 (+1)	8
	PCソフトウェア ・ツール等	表 A-8	11 (-1)	12 (-1)	13
	サービス	表 A-9	7 (-1)	8 (+2)	6
合計(のべ)			55 (-2)	57 (±0)	57

*表番号：合成製品一覧の対応表番号， **増減：前年度からの増減

音声合成製品の提供ベンダ数で見ると、「ミドルウェア PC 向けのその他」製品、「サーバ製品」、「PC ソフトウェア・ツール」、「サービス」などのカテゴリで、製品の提供を廃止するベンダがあり、ベンダ数が減少している。また、「ミドルウェア組み込み向けの規則音声合成」製品、「ミドルウェア PC 向けの規則音声合成」製品のカテゴリでは、平成24年度調査でいったん製品提供を中止したベンダが、平成25年度調査時に製品の提供を再開しており、ベンダ数が増える結果となった。

今回の音声合成の製品動向調査に製品情報を提供したベンダ数は、全カテゴリを合わせたのべ数で 55 社、複数のカテゴリに製品を提供しているベンダを考慮すると、音声合成製品を提供しているベンダは現在 31 社に上る。

○音声認識の製品動向調査

本年度調査にあたって、その製品分類・記載項目について検討し、昨年同様の分類ならびに記載項目で調査を実施することとした。具体的には、製品分類は音声認識製品を使用する立場から、①システム開発者向けとして、ミドルウェア（組込み向け）とミドルウェア（PC 向け）、②ソリューション利用者向けとして、CTI ソリューション、Web サービスソリューション、その他ソリューション、そして③エンドユーザ向けとして、応用製品とソフトウェア製品の、計 7 つに分類している。各分類における平成 22 年度～25 年度の製品件数とその推移を表 3 に示す。表 3 の注意点として、エンドユーザ向けの応用製品では、産業用もしくはオフィスにおける応用製品としており、調査の概要で述べた通り平成 23 年度からコンシューマ向け製品（カーナビゲーション、携帯電話、ゲームなど）を除外している。そのため平成 22 年度から平成 23 年度において、エンドユーザ向け応用製品の件数が大きく減少している。また、製品を提供しているベンダ数とその推移について表 4 に示す。

表 3 分類毎の音声認識製品の件数と推移

分類		表番号*	H25		H24	H23	H22
			件数	増減**	件数	件数	件数
①システム開発者向け 音声認識エンジン	ミドルウェア(組込み向け)	表 B-1	10	+2	8	8	11
	ミドルウェア(PC 向け)	表 B-2	7	+2	5	9	11
②ソリューション利用者 向け音声認識 ソリューション製品	CTI 分野	表 B-3	12	0	12	13	10
	Web サービス	表 B-4	4	+1	3	1	0
	その他	表 B-5	7	-2	9	6	7
③エンドユーザ向け 応用製品	応用製品	表 B-6	3	-2	5	5	20***
	応用ソフトウェア製品	表 B-7	13	0	13	15	21***
合計		-	56	+1	55	57	80***

*表番号：認識製品一覧の対応表番号， **増減：平成 24 年度からの増減

***：平成 22 年度まではコンシューマ製品（カーナビゲーション、携帯電話、ゲーム等）も含む

表 4 分類毎の音声認識製品のベンダ数とその推移

分類		表番号*	H25		H24	H23	H22
			件数	増減**	件数	件数	件数
①システム開発者向け 音声認識エンジン	ミドルウェア(組込み向け)	表 B-1	7	+1	6	5	7
	ミドルウェア(PC 向け)	表 B-2	5	+1	4	7	9
②ソリューション利用者 向け音声認識 ソリューション製品	CTI 分野	表 B-3	9	0	9	8	7
	Web サービス	表 B-4	4	+1	3	1	-****
	その他	表 B-5	4	-2	6	4	4
③エンドユーザ向け 応用製品	応用製品	表 B-6	2	-2	4	5	15***
	応用ソフトウェア製品	表 B-7	5	-1	6	10	8***
合計		-	36	-2	38	40	50***

*表番号：認識製品一覧の対応表番号， **増減：平成 24 年度からの増減

***：平成 22 年度まではコンシューマ製品（カーナビゲーション、携帯電話、ゲーム等）も含む

****：Web サービスの分類は平成 23 年度からのため、未計測

回答して頂いた 16 社に関する調査結果の詳細を表 B-1～B-7 に掲載する。

調査結果は、上記手法による本委員会からの依頼に対して回答のあった情報のみを掲載しているため、網羅性の視点からは完全な調査ではない。よって、周知であるにも関わらず調査結果に記載されていない製品が存在する。

また、昨年度は回答があったが今年度は回答が得られなかった機関についても、製品自体が継続して販売されていることが企業ホームページから確認できる場合でも、該当製品情報を掲載していない。

従って、本調査結果のみから、動向を判断することは困難であるが、表 3 からわかるここ数年の傾向を以下に示す。

- ・ システム開発者向けの音声認識エンジンについては、従来計数されていなかった製品が掲載されたため見かけ上製品数が増えているが、実質的には昨年度と同数である。また、バージョンアップされた製品は 3 製品である
- ・ ソリューション製品については、近年と同様の傾向であり、CTI 分野はほぼ横ばいだが、Web サービスは増加傾向にある。特にクラウドコンピューティングを利用したものが増えている
- ・ 全体の約 1 割にあたる 6 製品がバージョンアップされた。一方で新規追加された製品は 3 製品であり、安定的な推移であったと考えられる
- ・ ベンダ数については 2 社減少しており、この数年はわずかながら減少傾向がみられる

その他、本調査に含まれない委員会の独自調査からも、スマートフォン、携帯端末のアプリケーションが増加していることが本年度の大きな傾向であり、今後の動向が注目される。

なお、今後も、携帯端末上のアプリケーションや、サーバ上での音声認識利用サービスが増加することが予想され、製造企業、販売企業、サービス提供企業などが複雑化するとともに、販売製品という形態ではなく、有償・無償とにかかわらずエンドユーザに対するサービスとして提供される応用システムの増加も考えられる。次年度以降も、なるべく多くの情報を収集し、当該分野の動向を把握するとともに、関連製品開発者、導入検討者、エンドユーザに有益な情報を提供できるよう、調査方法について検討を続ける予定である。

○平成 25 年度の音声認識・合成技術による応用製品の全般的な動向

(音声合成)

2013 年 4 月には名古屋工業大学を中心としたプロジェクトから、音声合成ソフトウェア CeVIO(チェビオ)がリリースされた(本調査の製品リストには不掲載)。CeVIO はフリー版も配布されているソフトウェアで、音声合成エンジンをランタイムで利用するだけでなく、生成された音声波形(音声ファイル)をさまざまなコンテンツでのナレーションなどに利用可能としている点などが注目に値する。また CeVIO では、フェーダーによる感情のコントロールも可能であり、単にテキストを音声にして読み上げるだけにとどまらず、多彩な合成音声を出力する、というトレンドを作っている。

その他、本調査の調査対象ではないがエンドユーザにおける音声合成技術の応用傾向としては、2012 年 3 月に日本(語)でのサービスを開始した Siri、しゃべってコンシェルはその後も市場での知名度を定着させ、2012 年 10 月発売のモデル(2012 年冬モデル)からは「しゃべってキャラ」サービスを開始し、対応する声色のバリエーションを大幅に増加させた。これにより、平成 25 年度(2013 年度)には、音声対話アシスタントのアプリにおける音声合成技術は、単に「合成音声が使われる」というのみならず、「エンドユーザがさまざまな声色を好みに合わせて選べる」というトレンドを切り開いた。この背景には、いわゆる「カスタムボイス」の作成技術の低価格化、高性能化がある。

(音声認識)

音声認識については、製品動向調査の範囲外の応用として、エンドユーザ向けの音声対話アプリケーション、家電の対話的な操作などが増えている。また、Siri 等スマートフォンを利用したクラウド型音声対話サービスが普及し、音声対話が現実的になってきている。このようなサービスおよびそれを支える技術を活用する形で、Apple の CarPlay、トヨタの T-Connect などテレマティクスサービスにおける音声対話が再び脚光を浴びつつある。