

表2. 4. 1-2 音声合成ハードウェア製品(録音再生LSI)

分類	録音再生LSI	録音再生LSI	録音再生LSI	録音再生LSI	録音再生LSI	録音再生LSI
メーカー	沖電気工業	沖電気工業	沖電気工業	沖電気工業	沖電気工業	沖電気工業
品番名	ML2500B	ML2502	MSM9841	ML2302	MS87V1021	ML2308
符号化方式	analogストレージ	analogストレージ	4bitADPCM 4~8bitADPCM2 8,16bitPCM 8bit非線形PCM	4bitADPCM 2bit/4~8bitADPCM2 8,16bitPCM 8bit非線形PCM	4bitADPCM 4bitADPCM2 8bit非線形PCM(再生のみ)	$\mu$ -law(G.711準拠) 8/16bit PCM 8bit Oki非線形PCM 2~8bit Oki ADPCM2
ビットレート	Fsamに同じ	Fsamに同じ	16~512kbps	8~409kbps	16~42.4kbps	16~512kbps
接続メモリ	1Mbit AnalogFlash内蔵	128Kbit AnalogFlash内蔵	-	-	内蔵DRAM(2Mbit) 内蔵マスクROM(512Kbit)	-
最大アドレス	320	2	-	-	-	-
プロセス技術	CMOS	CMOS	CMOS	CMOS	CMOS	CMOS
マイクアンプ	ラインアンプ内蔵	内蔵	内蔵	内蔵	ラインアンプ内蔵	マイクアンプ × 2ch DRC機能付ラインアンプ × 2ch
ローパスフィルタ	内蔵	内蔵	内蔵	内蔵	内蔵	内蔵
A-Dコンバータ	8bit相当	8bit相当	14bit	14bit	14bit	1 bit $\Delta \Sigma$
D-Aコンバータ	8bit相当	8bit相当	14bit	14bit	14bit	1 bit $\Delta \Sigma$
原発振周波数	48.192MHz内蔵	48.192MHz内蔵	4.096MHz	4.096MHz	4.096MHz	24.576MHz
標準化周波数	4~6.4kHz	4~6.4kHz	4~32kHz	4~25.6kHz	4~10.6kHz	4~32kHz
電源電圧	2.7~3.3V	2.7~3.3V	2.7~5.5V	2.7~3.6V	2.7~3.6V	2.7~3.6V
消費電流(Max)						
動作時	45mA	40mA	30mA	20mA	15mA	40mA
スタンバイ時	10 $\mu$ A	10 $\mu$ A	10 $\mu$ A	10 $\mu$ A	40 $\mu$ A	20 $\mu$ A
パッケージ	32TSOP	30SSOP/CHIP	56QFP	64TQFP・71W-CSP	32TSOP	48QFN
価格	-	-	-	-	-	-
発売時期	-	-	-	-	-	-
備考		AGC 内蔵 SPアンプ 内蔵	1024bitバッファメモリ 音量調整機能 ステレオ出力	1024bitバッファメモリ SPアンプ内蔵 音量調整機能 早送り/巻戻し 早聞き/遅聞き	デレイ再生・リピート再生 早送り・巻戻し 過去録	1024bitバッファメモリ SPアンプ内蔵 ステレオ入出力 早送り/巻戻し 同時録音再生

表2.4.1-3 音声合成ハードウェア製品(再生専用LSI)

分類	再生専用LSI	再生専用LSI	再生専用LSI	再生専用LSI	再生専用LSI	再生専用LSI	再生専用LSI
メーカー	沖電気工業	沖電気工業	沖電気工業	沖電気工業	沖電気工業	沖電気工業	沖電気工業
品番名	MSM9802/ 03/05	ML2251/52/53/54/56	ML22Q54/Q58	ML2240	MSM9831/ML2201	ML2213/15	MSM9842
符号化方式	8bitPCM 8bit非線型PCM	2,4bitADPCM2 8,16bitPCM 8bit非線形PCM	2,4bitADPCM2 8,16bitPCM 8bit非線形PCM	2,4bitADPCM2 8,16bitPCM 8bit非線形PCM	8bit非線型PCM	4bitADPCM 8bitPCM・8bit非線型 PCM	4~8bitADPCM 8,16bitPCM 8bit非線型PCM
ビットレート	32~128kbps	8~768kbps	8~768kbps	8~768kbps	32~128kbps	16~128kbps	25.6~512kbps
接続メモリ	512K/1M/ 2Mbit内蔵	512K~6Mbit maskROM内蔵	4Mbit/8Mbit FlashROM内蔵	ROM外付(最大128Mbit)	384Kbit内蔵	1.5M/3Mbit内蔵	-
最大アドレス	63	256	256	256	31	247	-
プロセス技術	CMOS	CMOS	CMOS	CMOS	CMOS	CMOS	CMOS
マイクアンプ	-	-	-	-	-	-	-
ローパスフィルタ	内蔵	内蔵	内蔵	内蔵	内蔵	内蔵	内蔵
A-Dコンバータ	-	-	-	-	-	-	-
D-Aコンバータ	10bit	14bit	14bit	14bit	10bit	12bit	14bit
原発振周波数	4.096MHz	4.096MHz	4.096MHz	4.096MHz	4.096MHz	4.096MHz	4.096MHz
標本化周波数	4~16kHz	4~48kHz	4~48kHz	4~48kHz	4~16kHz	4~16kHz	6.4~32kHz
電源電圧	2.0~5.5V	2.7~3.6V 4.5~5.5V	Q54/2.7~3.6V Q58/2.7~5.5V	2.7~3.6V 4.5~5.5V	2.0~5.5V	2.4~5.5V	2.7~5.5V
消費電流(Max)							
動作時	16mA	35mA	35mA	20mA	3.9mA	4mA	30mA
スタンバイ時	10 $\mu$ A	15 $\mu$ A	55 $\mu$ A	15 $\mu$ A	10 $\mu$ A	10 $\mu$ A	10 $\mu$ A
パッケージ	18DIP・24SOP・ 30SSOP・CHIP	44QFP・ML2253/54のみ WCSP	44QFP	80TQFP	8SOP(M9831)/ 8SSOP(L2201)	14SSOP・24SOP(2213) 20SSOP・24SOP(2215)	56QFP
価格	-	-	-	-	-	-	-
発売時期	-	-	-	-	-	-	-
備考	編集ROM機能	2chミキシング 編集ROM機能	2chミキシング 編集ROM機能	4chミキシング 編集ROM機能	編集ROM機能 省スペース	メモリエ機能	1024bitバッファメモリ内蔵 音量調整機能 ステレオ出力

表2.4.1-4 音声合成ハードウェア製品(規則合成LSI)

分類		コーパスベース音声合成ボード	コーパスベース音声合成ボード
メーカー		エーアイ	エーアイ
商品名/品番名		HD AIvoice/CET1089EA	HD AIvoice/CET1089EB
入力形態		漢字仮名混じり文	漢字仮名混じり文
言語処理部	入力コード	シフトJIS	シフトJIS
	基本辞書	約15万語	約15万語
	ユーザ辞書	有	有
	処理性能	-	-
音声合成部	合成単位	可変長音素列	可変長音素列
	素片数	-	-
	パラメータ/方式	コーパスベース波形編集, 独自ピッチモデル	コーパスベース波形編集, 独自ピッチモデル
合成音声の種類	声・男女	自由に作成可(基本:女性)	自由に作成可(基本:女性)
	声の高さ	-	-
	発声速度	可変(0.7~1.3)	可変(0.7~1.3)
	その他	固定音声ファイル再生 句読点、改行等のポーズ設定可	固定音声ファイル再生 句読点、改行等のポーズ設定可
装置仕様 (ハード)	構成	SH3 160MHz RAM:8M/ROM:32M	SH3 160MHz RAM:8M/ROM:32M
	大きさ	幅:140mm 高さ:30mm 奥行き:90mm (突起物を除く)	幅:130mm 奥行き:80mm (突起物を除く)
	重量	約450g	-
	インタフェース	シリアル通信(RS-232C)	シリアル通信(RS-232C)
	電源, 電力	9V±10% 動作時 最大300mA	9V±10% 動作時 最大300mA
	出力	オーディオ出力 最大40mW(16Ω)30mW(32Ω)	オーディオ出力 最大40mW(16Ω)30mW(32Ω)
	その他	-	-
装置仕様 (ソフト)	ソフトウェア環境	-	-
	ハードウェア環境	-	-
	出力	-	-
価格		-	-
発表・発売時期		2005年6月	2005年6月
備考			

表2. 4. 1-5 音声合成ミドルウェア製品(録音再生ミドルウェア)

分類			
メーカー			
品番名			
符号化方式			
ビットレート			
プロセス技術			
マイクアンプ			
ローパスフィルタ			
A-Dコンバータ			
D-Aコンバータ			
原発振周波数			
標本化周波数			
電源電圧			
消費電流(typ)			
動作時			
スタンバイ時			
パッケージ			
価格			
発売時期			
備考			

本年度は、このカテゴリーの回答なし

表2. 4. 1-6(1) 音声合成ミドルウェア製品(規則合成ミドルウェア)

分類		ミドルウェア	音声合成ミドルウェア	音声合成ミドルウェア
メーカー		(株)アルカディア NECエレクトロニクス	開発:富士通・アニモ 販売:アニモ	開発:富士通・アニモ 販売:アニモ
商品名/品番		SpeeCAN-SDK TextToSpeechミドルウェア Ver.4	FineSpeech Basic	FineSpeech Ver2.1
入力形態		漢字かな混じり文 漢字仮名混じり文 韻律記号付き文字列	漢字仮名混じり文	漢字仮名混じり文
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS
	基本辞書	約150,000語	約8万語	15.9万語
	ユーザー辞書	有	有	有
	処理性能	英単語読み一部対応	—	—
音声合成部	合成単位	拡張音節単位	CV/VC	1ピッチ波形
	素片数	女声:約14万、男声:約6万	—	男女各約15,000種
	パラメータ/方式	波形編集方式(波形接続)	波形編集方式	1ピッチ波形編集, 独自ピッチモデル
				可変長音素列
合成音声の種類	声・男女	男・女	女声、男性	男声/女声
	声の高さ	固定	可変(21段階)	可変(5段階)
	発声速度	可変(1%~200%)	可変(21段階)	可変(10段階)
	その他	—	—	アクセントの強さ:4段階, 合成品質:3段階 抑揚4段階
装置仕様(ソフト)	ソフトウェア環境	WindowsXP,2000	ITRON, WindowsCE その他	Windows Me/2000/XP, PocketPC2002, 2003, Linux, ITRON Windows CE for Automotive WindowsAutomotive、他
	ハードウェア環境	CPU:Pentium III 500MHz 以上 メモリ:256MB以上 ディスク:2GB以上	VRシリーズマイコン上で動作 V850ファミリマイコン上で動作 DAコンバータ:16bit その他	CPU: Pentium150MHz以上 必要メモリ:16MB以上(128MB以上推奨) HDD:10MB以上の空き容量 その他組込製品(携帯電話等)
	出力	WAV、PCM、mulaw、G726形式での ファイル出力や直接Audio-outが可能。 サンプリング形式は22kHz/16bitなど から選択可能。	8kHz/11kHz/22kHz  16bit、モノラル	8kHz/11kHz/22kHz サンプリング  16ビットリニア、 $\mu$ law
価格	¥210,000(税込)~	別途ご相談	開発環境:50,000円、600,000円 ※再配布のロイヤリティは個別相談。	開発環境:200,000円、600,000円 ※再配布のロイヤリティは個別相談。
発表・発売時期	2002年10月	2002年2月	1999年9月	2003年12月
備考		従来の音声合成部に改良を加 データサイズを増加させることなく 音質の向上を実現しています。 カスタマイズ等につきましては、別 途ご相談させていただきます。	開発時のAPIとして、富士通独自 APIとMicrosoft Speech APIの 2種類を用意。	開発時のAPIとして、富士通独自API(FTTS-API) のほか、Microsoft Speech API(SAPI5.1)、ActiveX Control、SSMLに対応。  感情音声合成向けに「Emotional Option」 (100,000円)を別途提供。

表2. 4. 1-6(2) 音声合成ミドルウェア製品(規則合成ミドルウェア)

分類	ミドルウェア	ミドルウェア	ミドルウェア	テキスト音声変換ソフト
メーカー	スキャンソフト(株)	スキャンソフト(株)	スキャンソフト(株)	(株)ATR-Promotions
商品名/品番	ETI-Eloquence6.0	RealSpeak4.0	RealSpeakSolo4.0	音声合成SDK
入力形態	日本語:漢字カナ混じり文、韻律記号付カナ文字列、音素表記	日本語:漢字カナ混じり文、韻律記号付カナ文字列、音素表記	日本語:漢字カナ混じり文、韻律記号付カナ文字列、音素表記	Wizard Voice
	欧米語:アルファベット、音素表記	欧米語:アルファベット、音素表記	欧米語:アルファベット、音素表記	漢字仮名混じり文
言語処理部	入力コード	ASCII, UNICODE	ASCII, UNICODE	ASCII, UNICODE
	基本辞書	—	17万語	17万語
	ユーザ辞書	有	有	有
	処理性能	—	—	—
音声合成部	合成単位	—	音素、単語、文節	音素、単語、文節
	素片数	—	非公開	非公開
	パラメータ/方式	パラメータ合成	セグメント連結方式	セグメント連結方式
合成音声の種類	声・男女	男声、女声	男声、女声	男声、女声
	声の高さ	可変	変更不可	変更不可
	発声速度	可変	可変(10段階)	可変(10段階)
	その他	13ヶ国語(日、米、英、西、メキシコ、仏、加仏、独、伊、フィンランド、ブラジル、中、韓)	16ヶ国語(日、米、英、独、仏、伊、西、蘭、韓、北京、広東、メキシコ、ブラジル、ノルウェー、スウェー)	16ヶ国語(日、米、英、独、仏、伊、西、蘭、韓、北京、広東、メキシコ、ブラジル、ノルウェー、スウェー)
		カスタムボイス開発可能	カスタムボイス開発可能	カスタムボイス開発可能
装置仕様 (ソフト)	ソフトウェア環境	Windows, WinCE, Solaris, Linux, Embedded Linux, others under consideration	Windows 2000, Windows XP Professional, Windows 2003 Server Linux RedHat 7.2, Linux RedHat AS 2.1, Linux RedHat Enterprise WS 3.0, Linux RedHat Enterprise ES 3.0, Linux RedHat Enterprise AS 3.0 Solaris 8, Solaris 9(2005年予定)	Windows 98, Windows Me, Windows NT 4.0, Windows 2000, or Windows XP Microsoft Embedded Visual C++ 3.0 or later WinCE Version 3.0 or later The full speechbase versions are not included in WinCE
	ハードウェア環境		Pentium 4, XEON, minimum 1GHz, Single or multi-processor Recommended RAM: 1 GB RAM, minimum 512 MB Required disk space depending on language: up to 500 MB for language specific components	For windows, Intel Pentium-based PC (266 MHz) or higher 64Mb RAM . SoundBlaster 16-compatible audio device For winCE, Pocket PC 2003 Arm Compatible Processor (Strong Arm, Xscale, S3C2410)
	出力	8kHz, 11.025kHz	8kHz	11kHz, 16kHz, 22kHz
価格	下記へお問合せください	下記へお問合せください	下記へお問合せください	550000(税別途)
発表・発売時期	—	2005年1月	2004年秋	2002年8月
備考	埋め込み型システム用高音質テキスト合成エンジン 問合せ先: メール: Embedded_Japan@scansoft.com 電話: 03-5521-6011 FAX: 03-5521-6012	コンピュータ・テレフォニー用 高音質テキスト音声合成 問合せ先: メール: Network_Japan@scansoft.com 電話: 03-5521-6011 FAX: 03-5521-6012	埋め込み型システム用高音質テキスト音声合成 問合せ先: メール: Embedded_Japan@scansoft.com 電話: 03-5521-6011 FAX: 03-5521-6012	SDKの形式で提供しているので、Visual C/C++, Visual Basicでアプリケーション開発が可能。

\*記号の製品は、アンケート集計の都合上、昨年度のデータを使用

表2. 4. 1-6(3) 音声合成ミドルウェア製品(規則合成ミドルウェア)

分類	音声合成ミドルウェア		ミドルウェア	音声合成ミドルウェア	音声合成ミドルウェア
メーカー	NTTコミュニケーションズ/NTTアイ		旭化成(株)	エーアイ	エーアイ
商品名/品番名	PS2用音声合成ミドルウェア		VOStalk(ポストーク)	AIVoiceSDK Ver1.21	AIVoiceServer Ver1.31
入力形態	漢字仮名混じり文/韻律記号付カ		漢字仮名混じり文/韻律記号付き	漢字仮名混じり文	漢字仮名混じり文
言語処理部	入力コード	EUC	シフトJIS/EUC/UNICODE	シフトJIS	シフトJIS
	基本辞書	20万語	最大28万語(言語・用途に応じて語)	7万語~22万語	約15万語
	ユーザ辞書	有り	有り	有	有
	処理性能	頻出英単語読み対応	同形異読語, 表記ゆれ対応	同形異読語, 表記ゆれ, 英単語読	-
音声合成部	合成単位	音素(トライホン)	CV/VC	CV/VC	可変長音素列
	素片数	約6000個	言語による	-	-
	パラメータ/方式	波形編集方式(波形接続)	波形編集方式(波形重畳)	パラメータ合成方式	コーパスベース波形編集, 独自ピッチモデル
	パラメータ/方式	波形編集方式(波形接続)	波形編集方式(波形重畳)	パラメータ合成方式	コーパスベース波形編集, 独自ピッチモデル
合成音声の種類	声・男女	男声/女声/音声変換機能付き	男・女/言語	男声/女声	自由に作成可(基本:女性)
	声の高さ	可変(16階調)	可変	可変	可変(0.5~2.0)
	発声速度	可変(16階調)	可変	可変	可変(0.5~2.0)
	その他	-	-	スペクトルの特徴が変更可	エコー
	その他	-	-	スペクトルの特徴が変更可	エコー
装置仕様(ソフト)	ソフトウェア環境	-	-	Windows, WindowsCE, Linux, μC ※ANSI C準拠でOS非依存	Windows2000/XP/2003
	ハードウェア環境	DTL-T10000	32bit CPU(10~20MIPS程度) RAM:300KB~ ROM:350KB~ サンプリング周波数, 辞書サイズ等に	-	メモリ:256MB以上 空ディスク容量:500MB以上 *ディスクサイズは, 使用する音声データベースに依存
	出力	22KHz, 16bit, PCM, モノラル	-	8k/11k/16k/22k 16bit モノラル	16kHz, 16bitモノラルPCMデータ, 8bitモノラル μ-law PCMデータ
価格	開発キット: ¥300,000(本体価格) ランタイムライセンス: 個別相談	個別相談(ライセンス契約による)	個別相談	開発環境: 200,000~550,000 ※再配布のロイヤリティは個別相談。	開発環境: 800,000 ※再配布のロイヤリティは個別相談。
発表・発売時期	-	1999年5月	2004/4/1	2004年10月	2004年10月
備考	ゲームでの利用を想定した高速・高性能なテキスト音声合成ミドルウェア	<a href="http://web.canon.jp/technology/software/11.html">http://web.canon.jp/technology/software/11.html</a> (日本語) <a href="http://www.canon.com/technology/software/11.html">http://www.canon.com/technology/software/11.html</a> (英語)	低演算量・低メモリサイズでありながら、高品質な合成音声を實現、PCをはじめ、車載機器、携帯機器、ロボットなどの組み込み機器での利用が可能です。	個別の音声データベースを作成可 料金は別途相談	個別の音声データベースを作成可 料金は別途相談 インターフェースはXML形式 同時多重処理可

表2. 4. 1-6(4)音声合成ミドルウェア製品(規則合成ミドルウェア)

分類	音声合成ミドルウェア		音声合成ミドルウェア	音声合成ミドルウェア	音声合成ミドルウェア	音声合成ミドルウェア
メーカー	日本IBM		日本IBM	日本IBM	ペンタックス株式会社	ペンタックス株式会社
商品名/品番名	IBM Embedded ViaVoice, Multiplatform Edition		ViaVoice Runtime OEM	IBM WebSphere Voice Server V5.1	VoiceText 日本語	VoiceText 英語
入力形態	日本語:漢字仮名混じり文 欧米語:それぞれ各国語の文 音素表記		日本語:漢字仮名混じり文 その他言語:それぞれ各国語の文 音素表記	漢字仮名混じり文 SSML	漢字仮名混じり文 韻律記号付きかな文字列	英文
言語処理部	入力コード				日本語:シフトJIS	ASCII
	基本辞書	用途に応じて語数調整可能	用途に応じて語数調整可能	約14万語	約20万語	約12万語
	ユーザ辞書	有	有	有	有	有
音声合成部	処理性能	-	-	-	同形異読語、英単語読み	-
	合成単位	-	-	-	音素	音素
	素片数	-	-	-	-	-
合成音声の種類	パラメータ/方式	波形重畳方式	波形重畳方式	波形重畳方式	コーパスベース音声合成方式	コーパスベース音声合成方式
	声・男女	男声/女声	男声/女声	男声/女声	男声/女声	男声/女声
	声の高さ	可変(男声)	可変(男声)	可変	可変(50~200%)	可変(50~200%)
	発声速度			可変	可変(50~400%)	可変(50~400%)
装置仕様(ソフト)	ソフトウェア環境	様々なOSへ対応可能	様々なOSへ対応可能	* AIX 5. 3 * Red Hat Enterprise Linux WS/ES/AS (Intel 3.0 および SuSE SLES 8.0 版) * Microsoft Windows Server 2003	WindowsXP, 2000 Windows 2000 Server Windows 2003 Server RedHat Enterprise Linux AS/ES 3.0 Windows CE	WindowsXP, 2000 Windows 2000 Server Windows 2003 Server RedHat Enterprise Linux AS/ES 3.0 Windows CE
	ハードウェア環境	様々なプラットフォームへ対応可能	様々なプラットフォームへ対応可能	下記の URL を参照 <a href="http://www-306.ibm.com/software/pervasive/voice_server/">http://www-306.ibm.com/software/pervasive/voice_server/</a>	【PC版】 400MHz以上 必要メモリ:128MB以上 必要ハードディスク:760MB以上  【Windows CE版】 必要メモリ:14MB以上 必要ハードディスク:64MB,32MB	【PC版】 400MHz以上 必要メモリ:128MB以上 必要ハードディスク:650MB以上  【Windows CE版】 必要メモリ:16MB以上 必要ハードディスク:100MB-16MB
	出力	11kHz~22kHz, 16bit, モノラル	22kHz, 16bit, モノラル	8kHz, 16bit, モノラル	【PC版】 8kHz/16kHz モノラル 【Windows CE版】 11kHz モノラル	【PC版】 8kHz/16kHz モノラル 【Windows CE版】 11kHz モノラル
価格	お客様の使用する環境、システム構成により異なります。		お客様の使用する環境、システム構成により異なります。	要問合せ( <a href="http://www-6.ibm.com/jp/software/ecatalog/contactus/">http://www-6.ibm.com/jp/software/ecatalog/contactus/</a> )	ライセンス契約など条件で可変	ライセンス契約など条件で可変
発表・発売時	2004年11月		2002年9月	2005年11月	2004年1月~	2004年1月~
備考	モバイル・デバイス上で音声対応モバイル・ソリューションを作成するためのリソースをデベロッパーに提供するツールキットです。		様々なPC環境、モバイル・デバイス上で音声対応モバイル・ソリューションを作成するために必要な音声合成・認識のコンポーネントを提供します。	電話音声応答のアプリケーションを開発するための音声認識、および音声合成ソフトウェア。	<a href="http://voice.pentax.co.jp/">http://voice.pentax.co.jp/</a> TEL:03-3960-5246 E-mail:mmw@aoc.pentax.co.jp 限りなく自然でなめらかな音声合成です。	同左



表2. 4. 1-6(5) 音声合成ミドルウェア製品(規則合成ミドルウェア)

分類					
メーカー		ペンタックス株式会社	ペンタックス株式会社	(株)東芝	(株)東芝
商品名/品番名		VoiceText 中国語 音声合成ミドルウェア	VoiceText 韓国語 音声合成ミドルウェア	日本語音声合成ミドルウェア (TMW39-SYN-A)	日本語音声合成MW TTS型日本語音声合成MW (TMW39-TTS-A2)
入力形態		簡体字	ハングル文字	韻律記号付き文字列	日本テキスト(漢字仮名混じり文)/ 韻律記号付きカナ文字列
言語処理部	入力コード	GBK	KS_C_5601-1987	—	シフトJIS
	基本辞書	約7.9万語	約8.5万語	—	約14万語
	ユーザ辞書	有	有	—	設定可能
	処理性能	—	—	—	読みモード,速度,ピッチ等切替可能
音声合成部	合成単位	音素	音素	CV/VC	CV/VC
	素片数	—	—	302個	—
	パラメータ/ 方式	コーパスベース音声合成方式	コーパスベース音声合成方式	パラメータ合成方式(LPC)	パラメータ合成方式(LPC)
合成音声の種類	声・男女	女声	男声/女声	男声/女声	男声/女声
	声の高さ	可変(50~200%)	可変(50~200%)	可変(256段階)	可変(200段階)
	発声速度	可変(50~400%)	可変(50~400%)	可変(24段階)	可変(40段階)
	その他	音量(0~500%) 女声1話者	音量(0~500%) 女声1話者・男声1話者	声質9段階、抑揚512段階、 トーン10段階	音量(16段階) 声質(8段階)
装置仕様 (ソフト)	ソフトウェア環境	WindowsXP, 2000 Windows 2000 Server Windows 2003 Server RedHat Enterprise Linux AS/ES	WindowsXP, 2000 Windows 2000 Server Windows 2003 Server RedHat Enterprise Linux AS/ES Windows CE	—	組込開発環境はGHS OSには非依存
	ハードウェア環境	【PC版】 400MHz以上 必要メモリ:128MB以上 必要ハードディスク:650MB以上	【PC版】 400MHz以上 必要メモリ:128MB以上 必要ハードディスク:650MB以上 【Windows CE版】 必要メモリ:14MB以上 必要ハードディスク:12MB-100MB	東芝RISC(TX39シリーズ)	32bit CPU, Pentium
	出力	【PC版】 8kHz/16kHz モノラル	【PC版】 8kHz/16kHz モノラル 【Windows CE版】 11kHz モノラル	11.025kHz,16bit	モノラル 22.048kHz,11.025kHz 16bit
価格		ライセンス契約など条件で可変	ライセンス契約など条件で可変	個別相談	個別相談
発表・発売時期		2004年1月~	2004年1月~	2000年9月	2001年4月
備考		<a href="http://voice.pentax.co.jp/">http://voice.pentax.co.jp/</a> TEL:03-3960-5246 E-mail:mmw@aoc.pentax.co.jp 限りなく自然でなめらかな音声合成です。	同左		

表2. 4. 1-6(5) 音声合成ミドルウェア製品(規則合成ミドルウェア)

分類		音声合成ミドルウェア	音声合成ミドルウェア	音声合成ミドルウェア
メーカー		NTTアイティ	NTTアイティ	NTTアイティ
商品名/品番名		FutureVoice	FineVoice	Hipervoice
入力形態		漢字仮名混じり文/ 韻律記号付きカナ文字列/単語	漢字仮名混じり文/ 韻律記号付きカナ文字列/単語	漢字仮名混じり文/ 韻律記号付きカナ文字列/単語
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS
	基本辞書	50万語	20万語	20万語
	ユーザ辞書	有り	有り	有り
	処理性能	頻出英単語、人名読み対応	頻出英単語読み対応	頻出英単語読み対応
音声合成部	合成単位	-	その他(音素)	その他(音素)
	素片数	-	約9万(女声)、約2万(男声)	約6000個
	パラメータ/ 方式	コーパスベース音声合成方式	ハイブリッド合成方式	波形編集方式
	合成音声の種類	声・男女 声の高さ 発声速度	男声/女声 可変(100段階) 可変(100段階)	男声/女声 可変(100段階) 可変(100段階)
装置仕様 (ソフト)	ソフトウェア環境	Windows 2000/XP	Windows 95/98/Me/NT4.0/2000/XP	Windows 95/98/Me/NT4.0/2000/XP
	ハードウェア環境	CPU: Pentium III相当 1GHz以上推奨 メモリ: 100MB以上推奨 必要ハードディスク: 約1.3GB(8KHz)、 3GB以上(22KHz) Sound Blasterまたはその互換ボード	CPU: Pentium相当 300MHz以上推奨 メモリ: 64MB以上推奨 必要ハードディスク: 900MB以上(8KHz)、 2.3GB以上(22KHz) Sound Blasterまたはその互換ボード	CPU: Pentium相当 100MHz以上推奨 メモリ: 32MB以上推奨 必要ハードディスク: 30MB以上(8KHz)、 30MB以上(11KHz) Sound Blasterまたはその互換ボード
	出力	8kHz/22kHz、16bit、モノラル	8kHz/22kHz、16bit、モノラル	8kHz/11kHz、16bit、モノラル
	価格	個別相談	個別相談	個別相談
発表・発売時期				
備考		人間の声(肉声)の特徴を最大限に活かし、滑らかで自然な音声合成を実現する、新しいタイプの音声合成エンジンです。従来の規則合成方式の特徴であった機械的な音声ではなく、人の声に近い、聞きやすく親しみの持てる音声を作成することができます。辞書に登録されていない単語のアクセントも適切に付与可能です。ASPサービス「アナウンス代行サービス」もご提供しています。	音声品質を最大限に生かすNTT独自のハイブリッド合成方式によって肉声品質に迫る高品質合成を実現しました。日本語における音の連鎖を分析し、音素片接続の際の音質の劣化(低下)が最小となるよう工夫しています。	日本語テキスト文章を高品質な音声に変換するテキスト音声合成ソフトウェアです。漢字仮名混じり文を読みに変換し、自然なアクセントを付けて明瞭で滑らかな音声で読み上げます。

表2. 4. 1-7(1) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類		テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト
メーカー		クリエートシステム開発株式会社	クリエートシステム開発株式会社	クリエートシステム開発株式会社	クリエートシステム開発株式会社	クリエートシステム開発株式会社
商品名/品番名		Linux版 日本語音声合成ライブラリー	FreeBSD版 日本語音声合成ライブラリー	ドキュメントーカ for Windows Ver3.5	ドキュメントーカ for Mobile Ver2.0 視覚障害者向けPDAソフトウェア	ドキュメントーカ for Mac OSX Ver2.0
入力形態		漢字仮名混じり文/表音文字	漢字仮名混じり文/表音文字	漢字仮名混じり文/単語/英単語 /拡張MML言語による歌唱データ	漢字仮名混じり文/単語/英単語 /拡張MML言語による歌唱データ	漢字仮名混じり文/ 単語/英文
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS	シフトJIS	シフトJIS
	基本辞書	約14万語	約14万語	14万語	14万語	14万語
	ユーザー辞書	有	有	有	有	有
	処理性能	-	-	-	-	-
音声合成部	合成単位	1ピッチ波形編集	1ピッチ波形	1ピッチ波形	1ピッチ波形	可変長音素列
	素片数	男女各約15000種	男女各約15000種	男女各約15000種	男女各約15000種	約300文
	パラメータ/ 方式	1ピッチ波形編集 独自ピッチモデル	1ピッチ波形編集 独自ピッチモデル	1ピッチ波形編集 独自ピッチモデル	1ピッチ波形編集 独自ピッチモデル	コーパスベース波形編集 独自ピッチモデル
合成音声の種類	声・男女	男声/女声	男声/女声	男声/女声/ロボット音声など基本音声4種類 ユーザーからは「おじさん」「少年」など8種類の声として 設定可能	男声/女声	男声/女声/少年/少女/ロボット音声/英語 (SpeechManager)
	声の高さ	可変(5段階)	可変(5段階)	可変(5段階)	可変(5段階)	可変(5段階)
	発声速度	可変(10段階)	可変(10段階)	可変(10段階)	可変(10段階)	可変(10段階)
	その他	アクセント4段階	アクセント4段階	音量(10段階),トーン,エコー,波形伸縮	音量(10段階),トーン	音量(10段階),トーン,エコー,波形伸縮
装置仕様 (ソフト)	ソフトウェア環境	Linux	FreeBSD	Windows98/NT4.0/Me/2000/XP SAPI4	WindowsCE, HPC, PPC2000, PPC2002, PPC2003	Mac OSX V10.1以上
	ハードウェア環境	DOS_V機 容量,必要ボード等:サウンドボード	DOS_V機 容量,必要ボード等:サウンドボード	i486SX以上を搭載し,256色表示可能なWindowsが 稼動するPC。インストール時に約30MB以上の空き容量 のハードディスク。 最小搭載メモリ:32MB以上 サウンドカード:PCM録音再生,16ビット(8ビット), 11kHz,Mono,(Stereo),(MIDI)	Windows CE機 (HPC, PPC2000, PPC2002, PPC2003) インストール時に約4MB以上の空き容量が 必要	Mac OSX V10.1以上搭載  インストール時に約70MB以上の 空き容量のハードディスク, メモリ:10MB以上
	出力	16kHz,16bitモノラルPCMデータ, 8bitモノラルμ-law PCMデータ 11kHz,16bitモノラルPCMデータ, 8bitモノラルμ-law PCMデータ	16kHz,16bitモノラルPCMデータ, 8bitモノラルμ-law PCMデータ 11kHz,16bitモノラルPCMデータ, 8bitモノラルμ-law PCMデータ	11kHz, 16/8 bit, モノラル/ステレオ	8kHz, 16bit, モノラル	16kHz,16bit, モノラル
価格	CD-R版 9800円(税抜き) ベクターオンライン販売版: 4800円(税抜き)	CD-R版 9800円(税抜き) ベクターオンライン販売版: 4800円(税抜き)	7,800円(税抜き)	68,000円(税抜き)	7,600円(税抜き)	
発表・発売時期	1998年10月	2004年5月	2000年9月	2003年4月	2003年9月	
備考	富士通製音声合成エンジンのLinux版	富士通製音声合成エンジンのFreeBSD版	音声合成波形生成エンジンは富士通(株)が開発したものをWindows95/98/NT/XP用に移植。音声合成言語処理は独自の処理方式を採用。音声合成による歌唱が可能。 音声合成OCX、及びSAPI4が同梱されておりVisual Basic等でのアプリケーション開発が容易。	音声合成波形生成エンジンは富士通(株)が開発したものをWindowsCE用に移植。音声合成言語処理は独自の処理方式を採用。音声合成による歌唱が可能。 視覚障害者向けに特化	音声合成波形生成エンジンは富士通(株)が開発したFine Speech をMacintosh用に移植。音声合成言語処理は独自の処理方式を採用。英文テキストリグも内包。 音声認識も含むCarbon FrameworkとREALBasicの開発環境を提供している。	

表2. 4. 1-7(2) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類	方言ライブラリ	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	
メーカー	富士通ビー・エス・シー	富士通	PFU	クリエイティブシステム開発株式会社	クリエイティブシステム開発株式会社	クリエイティブシステム開発株式会社	
商品名/品番名	マルチメディア方言ライブラリ 「日本列島ことばの探検」 全国編 V1.0 監修・著:杉藤 美代子	VoiceScript2000 ランタイム	日本語音声合成機能 V 1.0	WindowsCE用音声合成SDK	ドキュメントーカ for Pocket PC	おもしろ替え歌 Ver2.0 for Mac OSX	
入力形態	単語リストや地図による選択	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文/単語/英単語	漢字仮名混じり文/単語/英単語	ひらがな/カタカナ	
言語処理部	入力コード	—	シフトJIS	EUC/シフトJIS	UNICODE	UNICODE	
	基本辞書	—	約16万語	12万5千語	14万語	14万語	
	ユーザ辞書	—	有	有	有	有	
	処理性能	—	—	—	—	—	
音声合成部	合成単位	音素	可変長音素列	1ピッチ波形(男女各約15000種)	1ピッチ波形	1ピッチ波形	
	素片数	約15000種	約300文	—	男女各約15000種	男女各約15000種	
合成音声の種類	パラメータ/方式	波形編集方式(波形接続)	コーパスベース波形編集, 独自ピッチモデル	1ピッチ波形編集 独自ピッチモデル	1ピッチ波形編集 独自ピッチモデル	1ピッチ波形編集 独自ピッチモデル	
	声・男女	男声	男声/女声	男声/女声	男声/女声	男声/女声	
	声の高さ	可変(音素ごとにマニュアルで設定)	可変(5段階)	可変(5段階)	可変(5段階)	可変(5段階)	
	発音速度	可変(音素ごとにマニュアルで設定)	可変(10段階)	可変(10段階)	可変(10段階)	可変(10段階)	
その他	—	アクセント4段階	アクセント4段階, 合成品質3段階	音量(10段階),トーン	音量(10段階),トーン	歌声合成,特殊MML	
装置仕様(ソフト)	ソフトウェア環境	Windows95&98/NT4.0	Windows2000	(日本語)Solaris 2.5.1, 2.6, (日本語)Solaris 7, 8, 9	WindowsCE日本語版	WindowsCE日本語版	
	ハードウェア環境	CPU:i486SX 66MHz以上 (Pentium90MHz以上推奨) 必要メモリ:16MB以上 (20MB以上推奨)	CPU:Pentium III 550以上 メモリ:128MB以上(4回線対応時) 音声処理ボード:富士通製FMVF-494Z又は インテル・ダイアロジックシステムズ社製互換ボード 音声FAX処理ボード:富士通製FMVF-493Z 又はインテル・ダイアロジックシステムズ社製互換ボード	sun4c, sun4m, sun4d, sun4u	Handheld PC, Pocket PC。 対応CPU:StrongARM (PXA,XScale),X86,SH インストール時に約4MB以上の空き容量, 16ビット,8kHz,Mono	Windows CE - Pocket PC 2002, (PXA,XScale) インストール時に約4MB以上の空き容量 16ビット,8kHz,Mono	Macintosh メモリ: 10 MB以上の空き ハードディスク容量:8MB以上の空き
	出力	22.05kHz, 16bit, モノラル	11.025kHz,16bit,モノラル	8kHz, 8bit μ law, モノラル	8kHz,16ビット,モノラル	8kHz,16ビット,モノラル	22kHz,16bit,モノラル,ステレオ
価格	19,800円	20万円～	125,000円	31,500円(税込)	ベクタープロレジサービス 2,000	ベクタープロレジサービス 2,500円	
発表・発売時期	1998年12月	1996年12月	1995年10月	2002年11月	2004年5月	2004年5月	
備考	時代と共に消えて行く方言。全国の懐かしい、また意外な方言音声を目と耳で体験する。 収録音声:(文部省研究助成「日本語音声」の一部)全国各地の単語・昔話「桃太郎」、琉球方言、アイヌ語の語り等 合成音声:富士通の音声合成手法で「かたつむり」「さつまいも」等の方言を一つずつ設定し作成。	地名辞書(全国31万地名)含む	C言語のAPIを用意しているため、APIを用いたアプリケーションの開発が可能。	音声合成波形生成エンジンは富士通(株)が開発したものをWindowsCE用に移植。 音声合成言語処理は独自の処理方式を採用。 <a href="http://www.createsystem.co.jp/download.html">http://www.createsystem.co.jp/download.html</a> よりダウンロード可能	音声合成波形生成エンジンは富士通(株)が開発したものをWindowsCE用に移植。 歌声合成エンジンを追加。 音声合成言語処理は独自の処理方式を採用。 <a href="http://www.createsystem.co.jp/download_PPC.html">http://www.createsystem.co.jp/download_PPC.html</a> よりダウンロード可能	音声合成波形生成エンジンは富士通(株)が開発したものを移植。 歌声合成エンジンを追加。 <a href="http://www.createsystem.co.jp/download_macosx.html">http://www.createsystem.co.jp/download_macosx.html</a> よりダウンロード可能	

表2. 4. 1-7(3) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類	テキスト音声変換ソフト		音声コンテンツ制作ツール		音声コンテンツ制作ツール	テレフォニーサーバ
メーカー	クリエイティブシステム開発株式会社		(株)アルカディア	NTTアイティ	NTTアイティ	NTTアイティ
商品名/品番名	ドキュメントーカ日本語音声合成エンジン/ドキュメントーカPlus		SpeeCAN-SFT	Visual Speech Creator II	Visual Speech Creator	ADVICE C3 (VXML対応 ボイスポータルプラットフォーム)
入力形態	漢字仮名混じり文/単語/英単語		漢字かな混じり文	漢字仮名混じり文/ 韻律記号付きカナ文字列/単語	漢字仮名混じり文/ 韻律記号付きカナ文字列/単語	
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS	シフトJIS	NTTアイティ「Hipervoice」、 「FineVoice」、「FutureVoice」 サーバ版に対応
	基本辞書	14万語	約150,000語	20万語	20万語	
	ユーザ辞書	有	有	有り	有り	
	処理性能	-	英単語読み一部対応	頻出英単語読み対応	頻出英単語読み対応	
音声合成部	合成単位	可変長音素列/1ビッチ波形	拡張音節単位	その他(音素)	その他(音素)	
	素片数	約300文/男女各約15000種	女声:約14万、男声:約6万	約9万(女声)、約2万(男声)	約6000個	
合成音声の種類	パラメータ/方式	コーパスベース波形編集/1ビッチ独自ピッチモデル	波形編集方式(波形接続)	ハイブリッド合成方式	波形編集方式	
	声・男女	男声/女声/少年/少女/ロボット音	男・女	男声/女声	男声/女声	
	声の高さ	音量(10段階),トーン	固定	可変(100段階)	可変(100段階)	
	発声速度	可変(10段階)	可変(1%~200%)	可変(100段階)	可変(100段階)	
その他	音量(10段階),トーン,歌声合成		-	-		
装置仕様(ソフト)	ソフトウェア環境	Windows98/NT4.0/Me/2000/XP SAPI4,SAPI5	WindowsXP,2000	Windows 95/98/Me/NT4.0/2000/XP	Windows 95/98/Me/NT4.0/2000/XP	Windows2000
	ハードウェア環境	i486SX以上を搭載し,256色表示可能なWindowsが稼動するPC。インストール時に約80MB以上の空き容量のハードディスク。最小搭載メモリ:128MB以上 サウンドカード:PCM録音再生,16ビット(8ビット),	CPU:Pentium III 500MHz 以上 メモリ:256MB以上 ディスク:2GB以上	CPU: Pentium相当 300MHz以上推奨 メモリ: 64MB以上推奨 必要ハードディスク: 900MB以上(8KHz) 2.3GB以上(22KHz) Sound Blasterまたはその互換ボード	CPU: Pentium相当 100MHz以上推奨 メモリ: 32MB以上推奨 必要ハードディスク: 30MB以上(8KHz)、 30MB以上(11KHz) Sound Blasterまたはその互換ボード	CPU:PentiumIII 1GHz以上推奨 メモリ:1GB以上推奨 ハードディスク容量:18GB以上 テレフォニーボード(NMS社製)
	出力	8-44kHz,16bit,モノラル,ステレオ	WAV、PCM、mulaw形式でのファイル出力や直接Audio-outが可能。サンプリング形式は22KHz/16bitなどから選択可能。	8kHz/22kHz, 16bit、モノラル	8kHz/11kHz, 16bit、モノラル	8kHz, 8bit、モノラル
価格	ベクタープロレジサービス 5,040円(税込み)		¥420,000(税込)~	398,000円(本体価格)	198,000円(本体価格)	個別相談
発表・発売時期	2005年9月		2002月10月	2001年10月	1999年8月	2001年8月
備考	音声合成波形生成エンジンは富士通(株)が開発したものを移植。音声合成言語処理は独自の処理方式を採用。 歌声合成エンジンを追加。 MS Visual Studio 2003,2005用クラスライブラリーも提供。 <a href="http://www.createsystem.co.jp/DTalkerSapi1.html">http://www.createsystem.co.jp/DTalkerSapi1.html</a>		GUIによるチューニング機能を装備(特許出願中)	従来の音声コンテンツ制作ツールの音声品質をNTT独自のハイブリッド合成方式により肉声品質に迫る高品質合成を実現しました。	合成音声のイントネーションを自由にデザインできる新しい音声コンテンツ制作ツールです。GUIによる操作で、音声の強さや長さを自在に編集できます。一度作成したイントネーションを自動的に再利用することも可能です。	VoiceXML対応、音声認識/音声合成搭載、IP電話対応(SIP)など高度な機能を多数備えた、ボイスポータルサービス構築のためのCTIプラットフォームです。

表2. 4. 1-7(4) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類		テキスト音声変換ソフト	テキスト音声変換ソフト	音声合成チューニングツール	PCアプリケーションソフト
メーカー		NTTデータ	三洋電機	エーアイ	株式会社言語工学研究所
商品名/品番名		ボイス君の テキストスピーチ2	ElegantalkVer3.0	AIVoiceTuningPro Ver1.1	韻律生成機能をもつ構文解析ソフト
入力形態		漢字仮名混じり文	漢字仮名混じり文/ 韻律記号付き仮名文字列	漢字仮名混じり文	漢字仮名混じり文
言語処理部	入力コード	シフトJIS	シフトJIS/UNI	シフトJIS	JIS/シフトJIS
	基本辞書	約27.7万語	27.7万語	約15万語	23万語
	ユーザ辞書	有	有	有	他に、構文解析のための200万行の付属語 無(登録可能)
	処理性能	同形異読語、英単語読み、 数字桁読みなど	同形異読語、英単語読み、 数字桁読み対応など	—	同形異読語、表記ゆれ、英単語読み対応あり
音声合成部	合成単位	VCVベース	VCVベース	1ピッチ波形編集	
	素片数	可変(～150MB)	可変(～150MB)	男女各約15000種	
	パラメータ/ 方式	波形編集方式	波形編集方式	1ピッチ波形編集 独自ピッチモデル	
合成音声の種類	声・男女	男声/女声	男声/女声2名/キャラクタ声	自由に作成可(基本:女性)	
	声の高さ	可変(251段階)	可変(251段階)	可変(0.5～2.0)	
	発声速度	可変(256段階)	可変(256段階)	可変(0.5～2.0)	
	その他		—	エコー	
装置仕様 (ソフト)	ソフトウェア環境	Windows98/Me/2000 Windows NT4.0/XP	Windows98/Me WindowsNT4.0/2000/XP日本語版	Windows2000/XP/2003	
	ハードウェア環境	Pentium166kHz以上、 メモリ容量32MB以上 ハードディスク 約105MB (推奨180MB以上)	Pentium166kHz以上、 メモリ容量32MB以上 ハードディスク 105MB以上 (音素片DBに依存)	メモリ:256MB以上 空ディスク容量:500MB以上 *ディスクサイズは、使用する 音声データベースに依存	
	出力	22/11/8kHz, 16 bit, モノラル	8/11.025/22.05kHz 16bit/PCMモノラル 8bit/μ-lawモノラル	16kHz,16bitモノラルPCMデータ, 8bitモノラルμ-law PCMデータ	
価格		8,925円/399,000円	ライセンス契約など条件で可変	550,000	
発表・発売時期		2004年2月	2004年2月	2004年10月	
備考		音声合成エンジン 「ElegantalkVer.3.0」 三洋電機機製	マルチプロセス/マルチスレッド対応 ActiveXコントロール	音素を自由に選択し、音声ファイルの 作成が可能	

\*  
\*記号の製品は、アンケート集計の都合上、昨年度のデータを使用

表2. 4. 1-7(5) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	
メーカー	NEC	NEC	NEC	NEC	
商品名/品番名	たび通(アメリカ旅行編)	HYPERVOICE-Light for TTS Premium	VoiceOperator	音声合成サポートキット	
入力形態	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文	
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS	
	基本辞書	約22万語	約20万語	約20万語	
	ユーザ辞書	無	有	有	
	処理性能	英単語読み対応など		英単語読み対応など	
音声合成部	合成単位	CV/VC	CV/VC	CV/VC	
	素片数	—	—	—	
	パラメータ/方式	波形編集方式	波形編集方式	波形編集方式	
合成音声の種類	声・男女	男声/女声	男声/女声	男声/女声	
	声の高さ	可変	可変	可変	
	発声速度	可変	可変	可変	
	その他	—	—	—	
装置仕様(ソフト)	ソフトウェア環境	Windows98, 98SE, Me, NT4.0, 2000, XP	Windows2000 Server	Windows2003 Server Windows2000Server(Service Pack 4), WindowsNTServer4.0(Service Pack 6以上)	
	ハードウェア環境	Pentium II 400MHz以上、 必要メモリ128MB以上 ハードディスク容量: 600MB以上	CPU:8回線以下 = 1.0GHz ~ 9回線以上 = 2.0GHz ~ メモリ:256MB(384MB) + 30MB × 回線数 + HVL 5MB × 回線数 HDD:50MB以上の空き	Pentium III 1GHz以上、 必要メモリ128MB以上 ハードディスク容量:最低150MB以上	Pentium III 1GHz以上、 必要メモリ256MB以上 ハードディスク容量:最低150MB以上
	出力	8kHz/11kHz/22kHz, 16bit, モノラル	8kHz, 8bit, モノラル μ-law	8kHz, 8bit, モノラル μ-law	8kHz, 8bit, モノラル μ-law
価格	29,800円	本体:800,000円 音声合成オプション(8回線ライセンス付き):300,000円 音声合成4回線ライセンス:150,000円	開発キット(25万円より), ランタイム(30万円より),他 音声合成サポートキット(本体5万、ライセンス7万より)	本体(5万円より)+ライセンス(7万円より)	
発表・発売時期	2001年8月	2002年5月	2003年10	2005年1月	
備考	旅行中に会おう場面に関連する日英音声認識・翻訳・読み上げ機能を搭載した旅行英会話支援ソフトウェアです。あなたの喋った日本語を認識して、それを英語に翻訳しさらに合成音声で結果を読み上げることができます。また、英和・和英辞典、会話の例文表示機能もあり、旅行英会話を支援してくれます。	アナログ電話回線を利用するIVRアプリケーション。FAX送信機能、録音機能、電話転送機能を有する。ver6.0にてマルチ音声合成サーバ対応。外部APのI/Fを提供しているため、外部AP連動も容易。またIVRの遠隔操作により、読み上げるテキストファイルを遠隔地から更新することが可能。	音声とFAXによる商品やサービスの案内、チケットや施設予約等、音声・FAX応答システムの構築と運用が可能。また、音声合成オプションにより、テキストファイルの内容や指定したテキスト(文字列)などの読み上げが可能(Text to Speech機能)。データベースから検索した内容などを確認する音声ガイドシステムにて、音声データの替わりにテキストの内容を読み上げるなど、IVR機能の利便性を高めることができる。	電話を利用したIVRソフトウェアに合成機能を追加できるAPIを提供。大規模なIVRシステムに対応し、より自然な合成音が可能(ニュースなどの自然な抑揚の読み上げが可能)。	

表2. 4. 1-7(6) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類	テキスト音声変換ソフト	ホームページ読み上げソフト	スクリーンリーダー	ホームページ作成ソフト	翻訳ソフト	
メーカー	日本IBM	日本IBM	日本IBM	日本IBM	日本IBM	
商品名/品番名	ProTALKER 97 Version 2.0	ホームページリーダー Windows 版 バージョン 3.01	JAWS for Windows (IBM Version) Version 3.7	ホームページビルダー バージョン6.5	インターネット翻訳の王様 バイリンガル Version4	
入力形態	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文	漢字仮名交じり文	
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS	シフトJIS	
	基本辞書	約14万語	約14万語	約14万語	約14万語	
	ユーザー辞書	有	有	有	有	
	処理性能	—	—	—	—	
音声合成部	合成単位	CV(環境依存型)	CV(環境依存型)	CV(環境依存型)	CV(環境依存型)	
	素片数	—	—	—	—	
	パラメータ/ 方式	波形編集方式	波形編集方式(波形重畳)	波形編集方式(波形重畳)	波形編集方式(波形重畳)	波形編集方式(波形重畳)
合成音声の 種類	声・男女	男声/女声	男声/女声	男声/女声	男声/女声	
	声の高さ	可変(10段階)	可変(10段階)	可変(10段階)	可変(10段階)	
	発声速度	可変(10段階)	可変(10段階)	可変(10段階)	可変(10段階)	
	その他	声の大きさ 可変(10段階) アクセントの強さ 可変(10段階)	—	—	—	
装置仕様 (ソフト)	ソフトウェア 環境	Windows 95, Windows NT	Windows 98, Millennium, 2000	Windows 98, Millennium, 2000	Windows 98, Millennium, 2000	
	ハードウェア 環境	CPU:Pentium 以上 必要メモリ:16MB以上 ハードディスク:20MB以上	CPU:Pentium300MHz以上 必要メモリ:64MB以上 必要ハードディスク:35MB以上	CPU:Pentium300MHz以上 必要メモリ:64MB以上 必要ハードディスク:130MB以上	CPU:Pentium以上 必要メモリ:32MB以上 必要ハードディスク:100MB以上	CPU:Pentium以上 必要メモリ:32MB以上 必要ハードディスク:80MB以上
	出力	22kHz/11kHz, 16bit/8bit, モノラル	22kHz/11kHz, 16/8bit, モノラル	22kHz/11kHz, 16/8bit, モノラル	22kHz/11kHz, 16/8bit, モノラル	22kHz/11kHz, 16/8bit, モノラル
価格	9,800円	15,000円	100,000円(98,Millennium版) 150,000円(2000版)	14,800円	9,800円	
発表・発売時期	1997年5月	2001年7月	発売中	2001年11月	発売中	
備考	MS Speech API準拠。開発者用キットあり。 簡易版を組み込んだ「ホームページリーダー」も97/10に発売	ホームページリーダーは目の不自由な方のインターネット・アクセスを可能にします。ホームページを合成音声で読み上げたり、電子メールの送受信を音声で確認しながら操作できます。 新バージョンではホームページへのアクセスがさらにパワフルになりました。	JAWS はWindowsの画面情報や入力を音声で読み上げるスクリーンリーダーです。 Windowsをはじめ各種アプリケーションの使用にいたるまで、キーボードだけの操作で画面情報や入力内容を読み上げさせることができます。	素材の作成から、ページ編集、サイト管理までを、これ1本で行なえるホームページ作成の統合ソフト。 日本語音声読み上げソフト「IBM ProTALKER97」に対応したしゃべるページが作成可能。簡単なスクリプトを取り込めば、ホームページの全体や一部をProTALKER97が読み上げてくれる	ホームページ、Eメールからテキストまで多目的に使える英日・日英翻訳ソフト。画面上の英単語をマウスでポイントすると英単語の意味が「王様くん」の吹き出しと音声で確認可能。また指定した原文や訳文も読み上げる。(ハントルの研究社の電子辞書とも連動)	



表2. 4. 1-7(7) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類	テキスト音声変換ソフト	テキスト音声変換ソフト	PCアプリケーションソフト	PCアプリケーションソフト	テキスト音声変換ソフト	
メーカー	日本IBM	ソニー	シャープ	株式会社システムソリューションセンターとちぎ	NTTデータ	
商品名/品番名	ViaVoice for Windows, Version 10.5 日本語版 (音声認識合成ソフト)	AIBOエンタテインメントプレーヤー ERS-7M3 (AIBOコントロール用ソフト)	メビウステロップ Ver.4	95Reader Ver6.0(XP Reader)	ボイス君のテキストスピーチ2(Standard)/(Professional)	
入力形態	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文	漢字仮名混じり文	
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS	シフトJIS	
	基本辞書	—	24万語	約20万語	10万語	
	ユーザ辞書	—	有	—	有	
	処理性能	—	—	—	漢字変換時の同音異義語の区別	
音声合成部	合成単位	—	CV/VC	VCVベース	—	
	素片数	—	—	—	—	
	パラメタ/方式	波形重畳方式	パラメタ合成・波形重畳方式	波形編集方式(波形重畳)ベース	—	
合成音声の種類	声・男女	男声/女声	男声/女声	男声/女声等(9話者)	男声/女声/ロボット	
	声の高さ	可変	可変	—	可変(9段階)	
	発声速度	可変	可変	—	可変(9段階)	
	その他	—	—	—	声の大きさ可変(20段階) アクセントの大きさ可変(10段階)	
装置仕様(ソフト)	ソフトウェア環境	Windows 2000/XP	Windows XP/2000	Windows Me, XP	Windows 98/98SE/2000/XP	
	ハードウェア環境	CPU:Pentium- II 333MHz 以上 必要メモリ:256MB以上 ハードディスク:600MB以上 22kHz, 16bit, モノラル	Pentium III 800MHz以上、 必要メモリ256M以上 ハードディスク 120MB以上	CPU:Pentium II 266MHz以上 ハードディスク:14MB以上(音声合成部)	CPU Pentium200MHz以上 必要メモリ 128MB以上 ハードディスク 300MB以上	Pentium 233MHz以上 メモリ容量64MB以上 ハードディスク空容量: (Standard) 約500MB以上、(Professional) 約500MB以上(標準インストール時)、約250MB以上(最小インストール時) ※すべての音声データベースをインストールするには1GB以上の空き容量が必要です。
	出力	22kHz, 16bit, モノラル	16kHz, 16 bit, モノラル	11kHz, 16bit	16bit/8bit ステレオ/モノラル	
価格	8,190円(Standard版 税込み)	194, 250円(税込)(AIBO本体を含む)	—	36,540円(税込)	8,925円(Standard) /399,000円(Professional)	
発表・発売時期	2005年9月	2005年10月	2002年1月	2004年4月	2004年2月	
備考	ViaVoice V10. 5に同梱の音声合成音声認識結果やメール文章などの読み上げが可能。	エンターテインメントロボットAIBO用。AIBOのリモートコントロール、音楽再生スケジュール読み上げなどのPCソフト、テキストを入力するとAIBOが読み上げ	シャープ製ノートPCの付属ソフトウェアインターネット等の情報をテロップ表示話し方(例:ささやき声)や面白いインターネットも設定可。	Windows版画面読み上げソフトウェア「95Reader」は、視覚障害者の方はもちろん、パソコンの情報を音声で知ることができるため、どなたでも使用可能なアクセシビリティ製品です。	音声合成エンジン「Elegantalk Ver.3.0」三洋電機㈱製 詳細URL <a href="http://www.voistage.com/products/vts02/">http://www.voistage.com/products/vts02/</a>	

表2. 4. 1-7(8) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類		テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト	テキスト音声変換ソフト
メーカー		東芝	東芝	東芝	東芝
商品名/品番名		EX-TREND 武蔵 Ver1 福井コンピュータ株式会社	ARCHITREND 21 Ver9 ARCHITREND Virtual House Ver4 ArchiMaster Ver6 福井コンピュータ株式会社	タオ 学習ソフト「天神」音声合成	TSOL 音声データによる検索システム 音声認識
入力形態		漢字かな混じり文	漢字かな混じり文	漢字かな混じり文	漢字かな混じり文
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS	シフトJIS
	基本辞書	約14万語	約14万語	約14万語	約14万語
	ユーザ辞書	有	有	有	有
	処理性能	同形異読語、英単語読み	同形異読語、英単語読み	同形異読語、英単語読み	同形異読語、英単語読み
音声合成部	合成単位	CV/VC	CV/VC	CV/VC	CV/VC
	素片数	284個	284個	284個	284個
	パラメータ/ 方式	パラメータ合成方式(LPC)	パラメータ合成方式(LPC)	パラメータ合成方式(LPC)	パラメータ合成方式(LPC)
	合成音声の種類	男声×2/女声×2	男声×2/女声×2	男声×2/女声×2	男声×2/女声×2
装置仕様 (ソフト)	声の高さ	可変(10段階)	可変(10段階)	可変(10段階)	可変(10段階)
	発声速度	可変(10段階)	可変(10段階)	可変(10段階)	可変(10段階)
	その他				
	ソフトウェア環境	2000/XP	2000/XP	2000/XP	2000/XP
ハードウェア環境					
出力	22.050kHz/16bit, ステレオ/モノラル	22.050kHz/16bit, ステレオ/モノラル	22.050kHz/16bit, ステレオ/モノラル	22.050kHz/16bit, ステレオ/モノラル	
価格					
発表・発売時期	2004年10月	2004年12月	2004年12月	2005/	
備考	建築設計CAD.CGシステム等のヘルプの読み上げに使用	建築設計CAD.CGシステム等のヘルプの読み上げに使用	学習ソフトの読み上げ		

表2. 4. 1-7(9) 音声合成ソフトウェア製品(音声規則合成アプリケーションソフト)

分類	組み込みシステム向け音声合成ミドルウェア	音声合成ミドルウェア	テキスト音声変換ソフト	PCアプリケーションソフト	PCバンドルソフト
メーカー	日立超LSIシステムズ	日立ケーイーシステムズ	日立ハイコス	日立製作所・日立ケーイーシステムズ	日立製作所・日立ケーイーシステムズ
商品名/品番名	RubyTalk SDK RubyTalk 評価キット	音声合成機能ソフトウェア開発キット Ver.3	流声 (音声アナウンスシステム)	心友	読みワザ
入力形態	漢字仮名交じり文 韻律記号付きかな文字列	漢字仮名交じり文 韻律記号付きかな文字列	漢字かな混じり文	漢字仮名混じり文	漢字仮名混じり文
言語処理部	入力コード	シフトJIS	シフトJIS	シフトJIS	シフトJIS
	基本辞書	非公開	非公開	-	-
	ユーザ辞書	有	有	登録可能	登録可能
	処理性能	同形異読語対応	-	-	-
音声合成部	合成単位	CV/V	CV/V	CV/V	CV/V
	素片数	-	-	-	-
	パラメータ/ 方式	波形編集方式(波形重畳)	波形編集方式(波形重畳)	波形編集方式(波形重畳)	波形編集方式(波形重畳)
合成音声の 種類	声・男女	女声	女声/男声	女声/男声	女声/男声
	声の高さ	可変(10段階)	可変(10段階)	可変(10段階)	固定
	発声速度	可変(10段階)	可変(10段階)	可変(10段階)	固定
	その他	音量(10段階)、抑揚強弱(10段階) エコー、高域強調	音量(10段階)、抑揚強弱(10段階) エコー、高域強調	アクセント(10段階)、ポーズ(時間指定)、音量(10段階)、抑揚(10段階)、	音量(3段階)
装置仕様 (ソフト)	ソフトウェア 環境	μ ITRON、T-Kernel。その他のOSに 関してはお問い合わせ ください	Windows2000/XP	非定型文編集:Windows 2000 Professional/Windows XP Professional 音声合成オンライン:Windows 2000	Microsoft Windows 98SE/Me/XP 日本 語版 Microsoft IME98/IME2000/IME2002
	ハードウェア 環境	CPU:SH3/SH3-DSP/SH4 ROM:10MB~48MB RAM:4.5MB	CPU:Pentium4 2GHz以上推奨 メモリ:256MB以上 ディスク:450MB以上の空き容量	CPU:Pentium 4 プロセッサ 2GHzクラス 以上推奨 メモリ:512MB以上 ディスク:1GB以上の空き容量 音源ボード、スピーカ LAN:アプリケーション、音声合成オン ライン、非定型文編集の通信用	CPU:Pentium 266MHz 以上 メモリ:32MB以上 ディスク:50MB以上の空き容量 解像度:1024×768ドット/800×600ドット 表示色:256色以上 音源ボード、スピーカ
	出力	22kHz, 16bit, モノラル	8/11/16/22kHz, 8/16bit, モノラル PCM/μ-Law	22KHz 16bit, モノラル	22kHz 16bit, モノラル
価格	弊社営業にお問い合わせください	開発環境300,000円~	ライセンス契約による	10,290円(税抜9,800円)	- (PCバンドル)
発表・発売時期	2005年10月	2004年12月	2005年11月	2004年12月	2005年4月
備考	日立超LSIシステムズは、音声合成ミドルウェアのトータル・ソリューションとして、組み込み機器で音声合成を使用するために必要な、ハードウェア/ソフトウェアを提供します。また、機器への組み込み、独自素片データの作成、コンサルテーションなどの技術サービスも提供致します。	肉声感が高くなめらかで自然な抑揚の高品位音声合成を実現。	アナウンス履歴保存、タフネス性強化等により高い信頼性が要求される情報制御システムへの適用が可能。	パソコンを使いたいけれど難しすぎて、と悩んでいるパソコン初心者やシニア年代の方々、パソコン画面の小さい文字が見えにくく目が疲れてしまうので困っている視力の弱い方のためのパソコン操作支援ソフトです。 <a href="http://www.hke.co.jp/products/shinyu/shinyu.htm">http://www.hke.co.jp/products/shinyu/shinyu.htm</a>	読みワザは、Internet Explorerで表示されるテキストや、クリップボードに保存されたテキストデータを読み上げるソフトウェアです。より肉声感のある自然な語り口で、漢字かな混じり文も適切に認識。また、BGMも同時に流すことができるので、まるでラジオを聞いているような感覚の楽しいソフトウェアです。

表2. 4. 1-8(1) 音声合成応用製品(組み込み機器他)

分類		その他	組み込み機器(カーナビ)	組み込み機器(カーナビ)	組み込み機器(カーナビ)
メーカー		日立製作所・日立ケーイーシステムズ	松下電器産業株式会社	松下電器産業株式会社	松下電器産業株式会社
商品名/品番名		伝の心	カーナビゲーション CN-HDS935MD	カーナビゲーション CN-HDS955MD	カーナビゲーション CN-HDS905D
入力形態		漢字仮名混じり文	JIS第一、第二水準漢字	JIS第一、第二水準漢字	JIS第一、第二水準漢字
言語処理部	入力コード	シフトJIS	-	-	-
	基本辞書	-	-	-	-
	ユーザー辞書	登録可能	-	-	-
	処理性能	-	-	-	-
音声合成部	合成単位	CV/VCV	CV/VCV	CV/VCV	CV/VCV
	素片数	-	-	-	-
	パラメータ/ 方式	波形編集方式(波形重畳)	波形重畳方式	波形重畳方式	波形重畳方式
合成音声の種類	声・男女	男/女	女声	女声	女声
	声の高さ	固定	固定	固定	固定
	発声速度	可変(3段階)	固定	固定	固定
	その他	-	-	-	-
装置仕様 (ハード)	構成	ノート型パソコン または デスクトップ型パソコン	64bit RISC CPU	64bit RISC CPU	64bit RISC CPU
	大きさ	-	本体:W178×H100×D184(mm)	本体:W178×H50×D160(mm)	本体:W178×H50×D160(mm)
	重量	-	3.4 kg	1.9 kg	1.9 kg
	インタフェース	-	-	-	-
	電源, 電力	AC100V	DC 12V 消費電流 3.5 A以下	DC 12V 消費電流 2.0 A以下	DC 12V 消費電流 2.0 A以下
	出力	ノート型:内蔵スピーカ/外部スピーカ/ ヘッドホン(イヤホン) デスクトップ型:外部スピーカ/ヘッドホン(イヤホン)	PCM 16bit 22.05kHz	PCM 16bit 22.05kHz	PCM 16bit 22.05kHz
	その他	-	-	-	-
装置仕様 (ソフト)	ソフトウェア環境	Windows XP 日本語版	-	-	-
	ハードウェア環境	DOS/Vパソコンで上記OSが動作する環境,  音源ボード、スピーカ	-	-	-
	出力	22kHz 16bit, モノラル	-	-	-
	その他	-	-	-	-
価格	50万円(非課税)	294,000円	320,250円	294,000円	
発表・発売時期	1997年12月	2005年	2005年	2005年	
備考	運動神経が侵される難病、筋萎縮性側索硬化症(ALS)患者がパソコンを利用して文書作成や、作成した文書の読み上げなどを行うことができる意志伝達装置。電子メールやホームページ閲覧、読書やテレビゲームなども楽しむことが可能。 <a href="http://www.hke.co.jp/products/dennosin/denindex.htm">http://www.hke.co.jp/products/dennosin/denindex.htm</a>	音声認識のトークバック、現在地読み上げ、FM文字多重放送読み上げ	音声認識のトークバック、現在地読み上げ、FM文字多重放送読み上げ	音声認識のトークバック、現在地読み上げ、FM文字多重放送読み上げ	

表2. 4. 1-8(2) 音声合成応用製品(組み込み機器他)

分類	組込機器	組込機器	組込機器	組み込み機器(コミュニケーションツール)
メーカー	キヤノン株式会社	キヤノン株式会社	キヤノン株式会社	廣済堂
商品名/品番名	ファクスホンSL75/SL55	ファクスホンPL95/PL90	imageRUNNER, Color imageRUNNER 音声ガイダンスキット(オプション)	SPコード専用読取機 スピーチオ
入力形態	音声操作ナビ, 着信読み上げ, L-mode電子メールの読み上げ, スケジュール読み上げ /漢字仮名混じり文	着信読み上げ /漢字仮名混じり文	操作・設定の音声ガイダンス(日本語, 英語)	SPコード(漢字仮名混じり文) 約800文字相当
言語処理部	入力コード	-	-	シフトJIS/ASCII
	基本辞書	-	-	約14万語
	ユーザ辞書	有り	有り	設定可能
	処理性能	同形異読語、表記ゆれ、英単語読み対応	同形異読語、表記ゆれ、英単語読み対応	読みモード、速度、ピッチ等切替可能
音声合成部	合成単位	CV/VC	CV/VC	CV/VC
	素片数	約300個	約300個	-
	パラメータ/ 方式	波形編集方式(波形重畳)	波形編集方式(波形重畳)	-
				パラメータ合成方式(LPC)
合成音声の種類	声・男女	男・女声	男・女声	男声/女声
	声の高さ	可変	可変	-
	発声速度	可変	可変	可変(200段階)
	その他			3段階 音量(16段階)
装置仕様 (ハード)	構成	L-modeホームファックス	L-modeホームファックス	imageRUNNER, Color imageRUNNERの オプション
	大きさ	SL75/SL55:344×248×281mm	PL95/PL90:363×300×223mm	-
	重量	SL75:約4.2kg、SL55:約3.9kg	PL95:約4.6kg、PL90:約4.4kg	-
	インタフェース	-	-	-
	電源、電力	AC100V 50/60Hz	AC100V 50/60Hz	-
	出力	SL75:8.9インチカラー液晶 SL55:5.1インチカラー液晶	PL95:タッチパネル式8.9インチカラー液晶 PL90:5.1インチカラー液晶	-
装置仕様 (ソフト)	その他	熱転写プリンタ	カラーインクジェットプリンタ	
	ソフトウェア環境	-	-	-
	ハードウェア環境	-	-	-
	出力	-	-	-
	その他	-	-	-
価格 発表・発売時期	オープン価格 2004年10月	オープン価格 2005年8月	オープン価格 2005年	128,000円
備考	<a href="http://cweb.canon.jp/faxphone/index.html">http://cweb.canon.jp/faxphone/index.html</a> <a href="http://web.canon.jp/technology/software/11.html">http://web.canon.jp/technology/software/11.html</a>	<a href="http://cweb.canon.jp/faxphone/index.html">http://cweb.canon.jp/faxphone/index.html</a> <a href="http://web.canon.jp/technology/software/11.html">http://web.canon.jp/technology/software/11.html</a>	<a href="http://web.canon.jp/technology/software/11.html">http://web.canon.jp/technology/software/11.html</a>	参照。スピーチオは、高齢者や目の不自由な方々に向けて開発された、SPコードの専用読取機です。SPコードを131万画素のCMOSイメージセンサで読み取り、記録された情報を音声で出力します(東芝日本語音声合成ミドルウェアTTS-A2を使用)。コードの位置合わせを行う台座が付いているほか、ボタン操作や音量の調節も、目の不自由な方々が操作しやすいように設計されています。