

Marktübersicht „Pick-by-Voice“ stellt Hersteller und Systeme gegenüber

Voice-Systeme erobern verschiedene Einsatzfelder in der Logistik

Die Idee, sprachgesteuerte Kommunikationssysteme im Lager einzusetzen, stammt ursprünglich aus den USA. In den letzten Jahren hat „Pick-by-Voice“ aber auch in Europa Erfolgsgeschichten geschrieben und stößt vor allem im Handel auf großes Interesse. „Logistik für Unternehmen“ fragte bei den Herstellern und Anbietern neben technischen Details auch aktuelle Entwicklungen, Preise und Erfahrungen ab.¹⁾

Hintergrundfragen:

Über die Angaben im Fragebogen hinaus, haben wir die Hersteller und Anbieter noch um Antworten zu den folgenden Fragen gebeten:

1. Welche Branchen setzen Pick-by-Voice verstärkt ein, und wo sehen Sie noch weiteres Marktwachstum?
2. Wo sehen Sie die Ursache für den verstärkten Einsatz von Voice-Systemen in den letzten beiden Jahren?
3. Wie werden sich die Preise entwickeln?
4. Wie hoch ist i.d.R. die Inbetriebnahmezeit inkl. Schulungszeit?
5. Welche Hauptstörpunkte gibt es bei Inbetriebnahmen?
6. Welche Entwicklungen stehen derzeit vor der Tür?

Sprachgesteuerte Systeme setzen sich im Bereich der Logistik und Produktion immer mehr neben beleg- oder terminalgestützten Prozessen durch. Typische Einsatzfelder sind

heute der Wareneingang, die Retourenabwicklung, die Kommissionierung oder die Qualitätskontrolle. Mit Headset und Terminal ausgestattet, kommuniziert der Anwender über Sprache

Bild 1 Mit freien Augen und Händen arbeiten Kommissionierer konzentrierter und schneller. Das führt zu mehr Produktivität, geringen Fehlerraten und hoher Arbeitssicherheit.



*Bild 2 Andreas Finken.
Bilder 2 und 5: topsystem*

mit dem Warehouse Management System (WMS) (**Bild 1**). Dazu bietet heute nahezu jedes WMS eine Schnittstelle zur Anbindung sprachbasierter Lösungen im Lager.

Nachdem wir bereits im Heft 11/2001, S. 6 bis 10 ausführlich sowohl über Funktionsweisen als auch Vor- und Nachteile der unterschiedlichen Spracherkennungssysteme berichteten, soll die vorliegende Marktübersicht die verschiedenen Hersteller und Systeme gegenüberstellen. Über die Angaben im Fragebogen hinaus, haben wir die Hersteller und Anbieter über Erfahrungen, Preise und aktuelle Entwicklungen befragt.

Wachstumspotenzial im Handel und bei Erweiterungen am größten

Dabei waren sich die befragten Unternehmen darin einig, dass das Einsatz-

potenzial von Voice-Systemen noch lange nicht ausgeschöpft ist. **Andreas Finken (Bild 2)**, Leiter Vertrieb und Marketing bei der topsystem Systemhaus GmbH, erwartet z.B. bei der Qualitätssicherung in der Produktion, in der Retourenbearbeitung und Warenverfolgung weiteres Marktwachstum.

Heidi Kühnert (Bild 3), Geschäftsführerin der proLogistik GmbH, und **Alexander Krutzke**, Produkt Marketing Manager bei Psion Teklogix GmbH, setzen weiterhin verstärkt auf den Handel, z.B. im Lebensmittelbereich, speziell im Tiefkühl-Lager, im Sanitärgrößhandel, im Textil-Bereich sowie in der Ersatzteillogistik.

Greg Tanner (Bild 4), Geschäftsführer von Vocollect Europe, sieht grundsätzlich in allen Branchen Wachstumspotenzial. Ein wichtiger Motor beim Neugeschäft sei zudem die Ergänzung bestehender Pick-by-Voice-Installationen um weitere Anwendungen, etwa



Bild 3 Heidi Kühnert.

Bild: proLogistik



Bild 4 Greg Tanner.

Bilder 1 und 4: Vocollect



um das Ent- und Beladen, die Einlagerung und Wiederauffüllung, das Cross-Docking oder die permanente Inventur.

Als Ursache für den bisherigen Erfolg sehen die Anbieter neben der Tatsache, dass diese Methode des Datenaustauschs nicht störend in den Arbeitsablauf im Lager eingreift und somit Verbesserungen hinsichtlich Produktivität, Fehlerraten und Arbeitssicherheit hervorbringt, verstärkt durch die mittlerweile ausgereifte Technik (Spracherkennung/Processorleistung/W-LAN/leichte Anbindung an vorgelagerte Systeme).

„Die Spracherkennung konnte deutlich verbessert werden und die Prozessorleistung in den Geräten ist gestiegen. Das führt zu schnelleren Antwortzeiten. Voice-Systeme sind heute keine Prototypen mehr“, erklärt **Kühnert**. Außerdem sind Voice-Systeme „relativ schnell zu installieren und amortisieren sich fast ausnahmslos innerhalb von drei bis zwölf Monaten“, ergänzt **Tanner**.

Bei der Entwicklung der Preise sind die Hersteller jedoch unterschiedlicher Auffassung. „Durch die immer größer werdenden Stückzahlen wird der Preis für die Hardware entsprechend nach unten gehen. Auf der anderen Seite werden neue Möglichkeiten wie RFID, Bluetooth integriert werden. Auch die Software wird bei gleichem Preis immer mehr Möglichkeiten bieten“, ist sich **Krutzke** sicher. **Tanner** stimmt dem nur bedingt zu: „Zwangsläufig entstehen durch wachsende Produktionszahlen Rationalisierungseffekte, die eventuell

Bild 5 Selbst bei einem hohen Geräuschpegel, wie hier verursacht durch Rollwagen, arbeiten sprachgesteuerte Kommunikationssysteme im Lager zuverlässig.

noch zu Preissenkungen führen.“ **Finken** erwartet eine langläufige Preisstabilität bis hin zu einer leichten Preissenkung. **Kühnert** geht dagegen von stabilen Preisen aus: „Es werden irgendwann schnellere Rechner zum gleichen Preis kommen.“

Dauer und Hauptstörfpunkte bei Inbetriebnahmen

Die Dauer der Inbetriebnahme hängt von vielen Variablen ab – natürlich von der Anzahl der zu installierenden Geräte, aber auch von der Zahl der Schichten, der vorhandenen Funknetz- und Serverstruktur sowie der Komplexität der jeweiligen Anwendung. Daher variiert die Inbetriebnahmezeiten von wenigen Tagen bis zu mehreren Wochen.

„Alle zur Erstinstallation notwendigen Schritte, inkl. des Anlegens von Aussprachemustern sowie der Einarbeitung von Lagerpersonal und Managern, sind i.d.R. innerhalb von ein bis eineinhalb Wochen abgeschlossen“, berichtet **Tanner**. **Kühnerts** Erfahrungen zeigen: „Die Inbetriebnahmezeit ist abhängig von der Anzahl der User und beträgt üb-

¹⁾ Interessenten können die Marktübersicht auch von der LOG-Website <http://www.technikwissen.de/logistik/aktuell/markt.asp> herunterladen.

licherweise drei bis fünf Tage vor Ort“.

„Das Sprachtraining dauert ca. 15 Minuten, dazu kommen noch zwei bis drei Stunden, um den kompletten Ablauf auf Sprachführung umzustellen“, beziffert *Krutzke* die Schulungs- und Inbetriebnahmezeiten.

„Wichtig im Rahmen der Inbetriebnahme ist die Schulung der Anwender im Umgang mit dem Sprachsystem. Anfänglich wurde dieser Aspekt für reibungslose Einführung des Sprachsystems unterschätzt“, weiß *Finken* aus Erfahrung. „Der Anwender muss sich erst daran gewöhnen von einer künstlichen Stimme geführt zu werden“, betont *Krutzke*.

Auch *Kühnert* schildert ähnliche Probleme: „Von unerfahrenen Anwendern werden in der Startphase häufig Erkennungsprobleme gemeldet, die ausschließlich aus Unkenntnis der zugelassenen Schlüsselwörter des aktuellen Dialogs resultieren.“

Aufgrund dieser Erfahrungen haben alle Anbieter die Abläufe bei Inbetriebnahmen optimiert. „Während der Inbe-

triebnahme ist der Versuch, alles extrem schnell zu implementieren, das größte Problem. Die Folge sind meist ein schlecht koordiniertes Vorgehen und Flüchtigkeitsfehler. Um das auszuschließen, nutzen *Vocollect* und seine Partner ein genau strukturiertes Implementierungsverfahren, das alle wichtigen Aspekte beinhaltet – von der Vorabplanung des Installationsprozesses bis zur Interaktion zwischen den Lagermitarbeitern, dem Management, den IT-Experten und dem Implementierungsteam“, berichtet *Tanner*.

Das bestätigt auch *Finken*: „Dank mehrjähriger Erfahrung mit der Einführung sprachbasierter Logistik-Anwendungen konnte von unserer Seite aus der Ablauf der Inbetriebnahme optimiert werden.“

Zukünftige Entwicklungen

Da die Anbieter von Voice-Systemen von ihren Kunden mit immer neuen Anforderungen konfrontiert werden,

wurde die Frage nach den zukünftigen Entwicklungen von den befragten Unternehmen sehr vielfältig beantwortet. So arbeitet *proLogistik* z.B. gerade daran, ihr System mehrsprachenfähig auch in anderen Ländern anzubieten. *Topsystem* sieht die nächste Entwicklungsstufe in der Integration von *BlueTooth* für die Kommunikation der einzelnen Hardware-Komponenten (PDA, Scanner etc.) untereinander, während *Psion Teklogix* den derzeitigen Trend darin sieht, sprachbasierte Systeme in *SAP* zu integrieren.

Vocollect kündigte hingegen an, noch in diesem Jahr weitere Headsets auf den Markt zu bringen. „Darüber hinaus optimieren wir zurzeit unsere *BlueStreak*-Spracherkennungs-Software für den Einsatz in besonders lauten Umgebungen. Zusammen mit einigen unserer Kunden und Partner arbeiten wir außerdem an einem *Transponder-Lese-Modul*, mit dem Anwender hören können, was ein *Transponder* ‚sagt‘“, so *Tanner* über die aktuellen Bestrebungen. **A.S.**

Marktübersicht: Pick-by-Voice-Systeme

Hersteller	Genesta, USA	proLogistik GmbH + Co KG	topsystem Systemhaus GmbH	Vocollect
Produktname	SyVox by Intermec	pL Voice	topSPEECH-Lydia-PDA	Talkman T2
weitere Anbieter/Partner (in D/A/CH)	Intermec Technologies, Düsseldorf (in Vorbereitung)	in.depend GmbH, Neuss Rodata AG, Dübendorf/CH	wird in D/A/CH von mehreren Partnern angeboten	IND Mobile Datensysteme, Psion Teklogix, Knapp, VoCognition
Hardware-Komponenten des Systems	Intermec 750 Terminal	x 86 kompatible eigengefertigte Rechnerbaugruppe inkl. Soundverarbeitung „on board“; Akkutechnologie LiION oder wahlweise NiMH	handelsübliche Hardware-Komponenten: PDA, Headset, Barcode-Scanner	Talkman T2-Terminal, verschiedene Headsets (normal, für kalte Umgebungen und "Behind-the-Neck"), Ladestationen für Terminals und Akkus, Akku-Analyse-Gerät, Gürtel mit Clip, Terminal-Hülle
Schnittstellen am Terminal (z. B. Anschluss für Scanner)	Barcode-Scanner, Touchscreen, Tastatur, RFID-Lesegerät	Headset, Barcode-Scanner, Tastatur und Monitor (für Service-Zwecke)	entsprechend des PDA-Modells, z. B. serielle Verbindung, Bluetooth	Je eine Schnittstelle für Headsets, für Trainings/Mithör-Geräte sowie für Barcode-Scanner und -Drucker
Gewicht des Terminals inkl. Akku	ca. 500 g	400g + 290g (Akku)	250 g	440 g mit Standard-Akku, 540 g mit Hochleistungsakku
Leistungsdauer des Akkus	8 h	4 bis 5 h	> 8 h, je nach System	Standard-Akku: min. 8 h, max. 10 bis 12 h. Der Hochleistungsakku hält noch länger.
Integrationsmöglichkeiten in bestehende Funksysteme	802.11b	Range Lan, 802.11b, 802.11g, 802.11a (Linux-Freigabe in ca. einem Monat)	entsprechend den PDA-Standards, z. Z. i.W. WLAN, 802.11b	802.11b
Höhe der übermittelten Datenrate	minimal	minimal, da reine Nutzdatenübertragung (100 – 500 Byte je Kommando)	11/54 Mbit/s	Abhängig von der jeweiligen Funk-Infrastruktur (2 MB FH/DSSS bis 11 MB)
Software-Komponenten des Systems	MS Pocket PC, SyVox Speech Client, SyVox Logistics Application (optional)	An alle LVS-Systeme anbindbar. Auf Wunsch kann ein Paket für die Kommissionierung – „pL Kom“ – mitgeliefert werden.	Spracherkennung, Sprachtrainer, Co-Pilot, topPICK-Voice	Standard: „VoiceClient“ inkl. „BlueStreak“-Spracherkennungs- und Text-to-Speech (TTS)-Sprachsynthese-Software, „Talkman Management Software“ Optional: „VoiceLink“ (Integrations-Tool), „VoiceManager“ (web-basiertes Verwaltungs-Tool)

Hersteller	Genesta, USA	pro Logistik GmbH + Co KG	topsystem Systemhaus GmbH	Vocollect
Arbeitsweise des Sprachverarbeitungssystems	IBM „ViaVoice“	Sprecherunabhängige phonem-basierte Spracherkennung direkt auf dem Client. Synthetische Sprachausgabe, daher minimaler Bandbreitenbedarf, hochskalierbar und flexibel.	Clientbasierte Spracherkennung, sprecherabhängig, sprachenunabhängig	Spracherkennung via sprecher-abhängigem Sprachmuster-Modell (basierend auf Hidden Markov Models); Sprachsynthese via TTS-Software (wandelt Textdaten in Sprache um) oder mit Hilfe der Record-and-Play-Methode (Sprachaufnahme und -wiedergabe)
Ist ein Standard-Wortschatz vorhanden? Wenn ja, wieviele Schlüsselwörter enthält dieser?	Ja, ca. 150 Wörter und Sätze. Individuelle Erweiterung auf Nachfrage.	Ja/z.Zt. ca. 70 Wörter.	Standardwortschatz umfasst ca. 20 Begriffe für die Spracheingabe. Der Wortschatz wird kundenspezifisch gestaltet.	Nein, Wortschatz ist abhängig von der Anwendung und kann bis 1 000 Wörter umfassen.
Muss bei einer Wortschatzerweiterung jeder Anwender geschult werden?	Nein.	Nein.	Ja, Nachtraining der neuen Worte.	Nein, bei einer Wortschatz-erweiterung muss nicht der Anwender, sondern das System trainiert werden: Von jedem Wort und für jeden Nutzer wird auf dem Terminal ein Sprachmuster bzw. Aussprachemuster angelegt, das anschließend von der Software wiedererkannt wird.
Ist ein separater Server für die Sprachverarbeitung und/oder -steuerung notwendig?	Nein, die Sprachverarbeitung erfolgt auf dem mobilen Endgerät	Nein.	Nein.	Nein, alle Spracherkennungs- und Sprachsynthese-Prozesse erfolgen innerhalb des Talkman-T2-Terminals.
Fremdsprachenunterstützung	Alle Sprachen über IBM "ViaVoice" abrufbar.	Ab Oktober 2004 lieferbar.	topSPEECH-Lydia-PDA ist sprachenunabhängig. Somit sind beliebige Fremdsprachen möglich.	Ja, aufgrund der sprecherabhängigen Gestaltung „verstehen“ die Spracherkennungssoftware jede Sprache, jeden Dialekt und jeden Akzent des Sprechers. Die TTS-Software zur Umwandlung von Textdaten in Laute ist in zahlreichen Sprachen erhältlich. Wenn in einer Sprache keine TTS-Software existiert, kann auf die Record-and-Play-Methode zurückgegriffen werden.
durchschnittliche Einarbeitungszeit pro Anwender in der Kommissionierung	30 min	4 h	ca. 20 min in Abhängigkeit von der Anwendung	15 bis 20 min für das Anlegen der Aussprachemuster, anschließend je nach Anwendung weitere 30 bis 90 min für die Eingewöhnung.
garantierte Spracherkennung bis zu einem Geräuschpegel von	ca. 85 dbA	Es liegen keine Messungen vor.	abhängig von der Anwendung und der Lagerumgebung, getestet bis 92 dbA	Getestet bis 93 dbA ohne ernst zu nehmende Störung der Spracherkennung
Einsatzmöglichkeiten	Kommissionierung, Auszählung, Einlagerung, Inspektion	Kommissionierung, Wareneingang, Transportsteuerung, Warenausgangskontrolle, Nachschub, QS-Kontrolle	Kleinteilkommissionierung, Getränkegroßhandel, Tiefkühlbereich, Frischekommissionierung, Wareneinlagerung, Warenausgang, Nachschub, Containerverwaltung	Kommissionierung, Annahme, Einlagerung, Wiederbefüllung, Cross-Docking, Beladung, Lieferung und permanente Inventur – die Anwendungsmöglichkeiten sprachbasierter Lösungen sind nahezu endlos.
Systempreis bei einer Größe von 15 Anwendern / 50 Anwendern	ca. 4 500 Dollar pro Kommissionierer	15 Anwender = 75 000 Euro (Komplett-Preis) 50 Anwender = 250 000 Euro (Komplett-Preis)	15 Anwender = 75 000 Euro (ab ca. 30 000) 50 Anwender = 185 000 Euro (ab ca. 90 000), je nach Leistungsumfang	Preis-Informationen über Vertriebspartner
Bisher größte realisierte Anwendung an einem Standort (Anzahl parallel laufender Geräte)	keine Angaben	62	80	rd. 300
Referenzen (Top 3)		Rewe, Lehrte; Bührmann Weine; Rewe, Wiesloch/ Norderstedt	Kaufhof Warenhaus AG, Lekkerland-Tobaccoland, Bauer AG, Trinks GmbH	Wal-Mart, Carrefour, Globus

Adressen:

Genesta Managing Partner
1850 Interstate 30
Rockwall, Texas 75087 USA

proLogistik GmbH + Co KG
Fallgatter 1
D-44369 Dortmund

topsystem Systemhaus
GmbH
Europark A2 Adenauer-
straße 20
D-52146 Würselen

Vocollect Europe
Gemini House
Mercury Park
Wycombe Lane, Wooburn Green
Bucks, HP10 0HH UK