Frischer Wind

Automatisierte Sprachanwendungen sind ein Wachstumsmotor im Telefongeschäft

Christoph Sieberz

Das klassische Internet bekommt Konkurrenz: Der Zugriff auf Anwendungen und Transaktionen sowie das Navigieren durch Informations- und Entertainment-Inhalte wird künftig zunehmend über Telefon und Handy ablaufen. Sprachverarbeitende Applikationen der neuen Generation nutzen das Internet als Plattform und Sprache als Mittel für die Interaktion und das (Mobil)Telefon als Medium. Weil sie durch ihren hohen Benutzerkomfort und die einfache Handhabung über Ohr und Stimme wertvolle Services erschließen, wird den Sprachanwendungen eine glänzende Zukunft vorausgesagt. Dabei werden die Telefongesellschaften von dem neuen Wachstumsmarkt besonders profitieren. Steht dem Voice Business in Zukunft gar ein ähnlicher Wachstumsschub bevor wie vor kurzem dem Markt für Internet-Datendienste?

Christoph Sieberz ist Vorstand Marketing bei der VoiceObjects AG in Bergisch Gladbach

Vieles spricht dafür, daß dem Voice Business die Zukunft gehört. In einer gemeinsam im deutschsprachigen Raum durchgeführten Studie kommen die Beratungsgesellschaften Heitzig Consult und Strateco („Sprachapplikationen auf dem Prüfstand", Oktober 2003) zu dem Schluß, daß bereits im Jahr 2005 jeder Vierte eine oder mehrere Sprachapplikationen regelmäßig nutzen wird. Auf deutschsprachigen Sprachdialogplattformen laufen, so die Studie, schon heute täglich 1,8 Mio. Vorgänge ab. Nach einer relativ aktuellen Untersuchung von Datamonitor („Voice Portals and Applications", Juni 2002) wird der Weltmarkt für sprachverarbeitende Lösungen – wohlgemerkt ohne die damit generierten Umsätze – von heute 656 Mio. $ auf 3,1 Mrd. $ im Jahr 2008 wachsen.
Diese Wachstumsprognosen könnten sogar noch übertroffen werden, wenn der Markt eine gewisse Eigendynamik entwickelt. Und das scheint zumindest nicht unwahrscheinlich, wenn man ihn mit dem Internet vergleicht: Sprachportale werden künftig wie Websites zu einem Voice Web verlinkt. Wie beim WWW werden mit der Menge und der Attraktivität der Inhalte die Nutzerzahlen steigen. Und mit den Nutzern wiederum wachsen die Zahl der Voice-Web-Auftritte und das Inhalteangebot.

Das Thema in Kürze

Sprachanwendungen der neuen Generation haben ein enormes Marktpotential. Unternehmen und Behörden profitieren vom neuen Kommunikations- und Absatzkanal, die Anrufer vom hohen Komfort des immer verfügbaren Mediums. Vom zu erwartenden erhöhten Daten- und Sprachverkehr werden vor allem die Telcos profitieren. Voraussetzungen für den Erfolg sind wirtschaftliche Systeme und die Entwicklung großer, massentauglicher Anwendungen.



Mit der Unterstützung durch virtuelle Kollegen können sich Call-Center-Agents mehr um individuelle Anfragen kümmern

(Foto: Siemens)

Treibende Kräfte

Das ist aber nicht die einzige Parallele zum WWW: Wie beim Internet treiben wirtschaftliche Erwägungen wie Kostensenkung, Kundenbindung, Produktivitätssteigerung und nicht zuletzt das Bestreben, neue Einnahmequellen zu erschließen, den Markt der Sprachanwendungen voran. Die entsprechenden Trends heißen Customer Self Service, innovative Geschäftsmodelle und mobiles Arbeiten.
Ein starker Markttreiber ist auch der Bedarf an höherer Bedienfreundlichkeit bei der Interaktion mit den immer leistungsfähigeren, aber zunehmend komplexer werdenden Anwendungen.
Zudem entwickelt sich der Markt nicht zuletzt mit dem Fortschreiten der Technik und den fallenden Preisen für die Infrastruktur. Nicht zu unterschätzen sind bei der Einschätzung des künftigen Wachstums auch die Marketingbemühungen der Marktteilnehmer.

Telcos profitieren

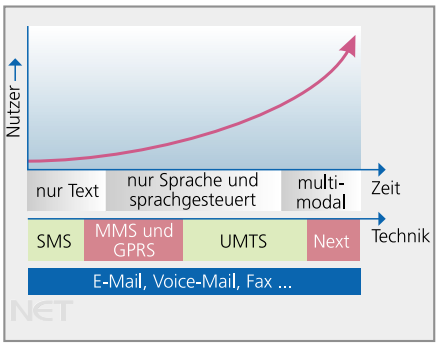
Die wichtigsten Player auf diesem Markt und zugleich diejenigen, die von den Sprachanwendungen am meisten profitieren werden, sind die Carrier. Mehr noch, aus verschiedenen Gründen sind sowohl die Fest- als auch die Mobilfunknetzbetreiber geradezu auf eine breite Nutzung von Sprachanwendungen angewiesen. Auf der Agenda vieler Carrier steht das Thema zur Zeit beim Business Development ganz oben. Bei der Deutschen Telekom befassen sich beispielsweise gleich alle drei Bereiche – T-Com, T-Mobile und T-Systems – aktuell sehr intensiv mit konkreten Voice-Projekten.

Die TK-Unternehmen suchen nach neuen Einnahmequellen, weil die Einnahmen in ihrem Kerngeschäft, den klassischen Telefongesprächen, stagnieren oder sogar rückläufig sind. Sprachbasierte Services sollen Abhilfe schaffen, denn jeder Dialog mit einer Sprachanwendung generiert Airtime und erhöht den viel beschworenen ARPU (Average Revenue Per User). Attraktive Sprachinhalte und hochwertige, sprachgesteuerte Dienste bilden so die Basis für das zukünftige Wachstum im Marktsegment Sprachverkehr.

Sprache treibt Daten

Zudem setzt vor allem die Mobilfunkbranche auf die Einnahmen aus Datenservices. Sollen sich die milliardenschweren Investitionen in UMTS-Lizenzen jemals annähernd bezahlt machen, müssen die Einnahmen aus Datendiensten signifikant steigen. Und auch wenn es zunächst paradox erscheint: Um den Markt der mobilen Datendienste in Gang zu bringen, werden die Sprachtechnik und darauf basierende Anwendungen eine bedeutende Rolle spielen.

Mit der Möglichkeit, mit Hilfe von VUI (Voice User Interface) komplexe Anwendungen über ein einfaches Telefon zu steuern, führen die Carrier die Benutzer an telefonbasierte Datendienste heran und eröffnen damit einen Evolutionspfad für die Endgeräte der nächsten Generation. Denn in der Zukunft werden die Grenzen zwischen Sprach- und Datenanwendungen zunehmend fließend.



Die Sprachsteuerung wird das Wachstum der Datendienste vorantreiben

Zu diesem Schluß kommt jedenfalls Forrester Research in der Marktanalyse „Mobile Applications That Drive Revenue". In ihr wird die Rolle der Sprachapplikationen als Trojanisches Pferd beschrieben, um lukrative Datenservices quasi durch die Hintertür einzuführen. Auch der Begriff „Multimodalität" spielt in diesem Bericht eine große Rolle.

Multimodalität und Multimedia

Was verbirgt sich dahinter? Multimodalität bedeutet schlicht, daß die Interaktion mit einer Applikation, abhängig von der augenblicklichen Situation des Benutzers, über verschiedene Schnittstellen erfolgen wird. So kommt als Hilfsmittel für die Eingabe wahlweise die Stimme oder die Hand (über Tastatur, Stift oder Touchscreen) zum Einsatz. Und die Ausgabe der Daten erfolgt – ebenso frei wählbar – entweder visuell über Displays oder akustisch durch Sprachausgabe. Beim Autofahren bietet sich beispielsweise der Sprachdialog als „freihändiger" Modus an. Bahnfahrer werden eher die diskreten, stummen Interaktionsmöglichkeiten bevorzugen.

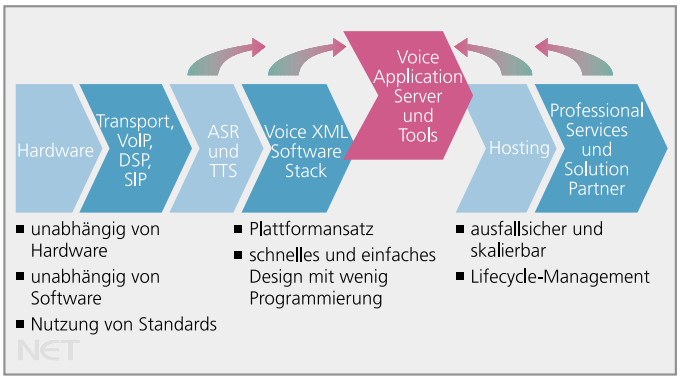
Welches der jeweils komfortabelste Interaktionsmodus ist, hängt auch von der Größe des benutzten Endgerätes ab. Bei mobilen Geräten wie GPRS- oder UMTS-fähigen Mobiltelefonen, Smartphones oder PDAs ist ein Trend zur Miniaturisierung zu erkennen, der der Benutzerfreundlichkeit allerdings entgegen läuft. Immer leistungsfähigere, dadurch aber schneller und komplizierter werdende Datenservices lassen sich zudem nur noch mit Einschränkungen über die Tastatur bzw. den Stylus steuern.

Mit Sprachsteuerung könnten auch diese Geräte im Taschenformat dagegen problemlos genutzt werden. Das gilt nicht nur für die mobilen Business-Services, sondern auch für die Handy-basierten Multimediaanwendungen wie Spiele.

Zum Konzept der Multimodalität gehört auch, daß die verschiedenen Ein- und Ausgabemodi stets parallel verfügbar sind und der Nutzer in der Interaktion jederzeit zwischen Auge und Ohr, Hand und Stimme wechseln kann. Dazu müssen die Endgerätehersteller aber noch einige Hausaufgaben erledigen. Wichtig bei UMTS ist beispielsweise eine Zusammenführung des Sprach- und Datenkanals.

Multimodalität könnte sich als eine der Schlüsseltechniken für die nächste Mobilfunkgeneration erweisen. Kein Wunder also, daß die Anbieter die entsprechenden Techniken mit Nachdruck vorantreiben.

Wegen ihrer wichtigen Rolle in der Wertschöpfungskette müssen Voice Application Server gut auf die verschiedenen Marktteilnehmer abgestimmt sein



Selbstbedienung über das Telefon

Kommen wir zum nächsten wichtigen Markttreiber. Unternehmen und Behörden suchen gerade in wirtschaftlich schwierigen Zeiten nach Wegen, um die Kosten im Kundenkontakt zu senken und gleichzeitig Kunden mit besserer Servicequalität zu binden.
Nachdem das Internet einen neuen Kommunikationskanal zwischen Unternehmen und Kunden eröffnete, der in der Regel allerdings nur von einem Teil der Kunden gut angenommen wurde, besinnen sich viele Unternehmen heute wieder auf das Telefon als wichtigstes Instrument im Kundenkontakt. Schließlich verfügt nahezu jeder über mindestens ein Telefon, was den potentiellen Nutzerkreis gegenüber dem Internet deutlich erhöht.
Aktuelle Studien belegen zudem, daß trotz des Informationsangebotes im Internet rund 70 % der Kommunikation mit dem Kunden noch immer aus der telefonischen Abfrage einfacher Informationen besteht, wie etwa bei Banken die Abfrage des Kontostands. Kein Wunder also, daß das Telefon als mit Abstand bedeutendster Kommunikationskanal wieder in den Mittelpunkt des Interesses und der technischen Entwicklung rückt.
Was liegt also näher, als die Call Center und die Telefonzentralen als zentrale Drehscheiben des Telefonkontakts durch automatisierte Sprachanwendungen zu ergänzen und somit zu entlasten? Die Mitarbeiter könnten sich dann verstärkt um beratungsintensive Gespräche mit Kunden, die individuelle Wünsche haben, kümmern. Mit Customer Self Services über das Telefon können Unternehmen und Behörden den Kundenkontakt zugleich rationalisieren und verbessern.
Einer der Vorreiter ist die Bankenbranche mit ihren automatisierten Telefon-Banking-Lösungen der neuen Generation. Verglichen mit der zurückliegenden Generation der Marke „Blechstimme" zeichnen sich moderne Sprachanwendungen durch einen hohen Benutzerkomfort aus und ermöglichen bereits personalisierte Services und Inhalte. Auch Transaktionen wie Buchungen und Aktiengeschäfte lassen sich damit zu jeder Tages- und Nachtzeit realisieren. Damit könnten nach einer Einschätzung der Giga Information Group bis zu 30 % der Call-Center-Kosten eingespart werden.
Weitere Branchen mit vergleichbar intensivem Kundenkontakt und größtenteils einfachen, standardisierbaren Anfragen und Prozessen sind beispielsweise Versicherungen, die Luftfahrt- und Bahnunternehmen, aber auch die Fernseh- und Hörfunkanstalten mit ihren Telefon-Votings und Gewinnspielen für ein Massenpublikum.

Geschäfte mit dem Telefon als Massenmedium

Womit wir bei den Medien angekommen wären. Nach einer Umfrage des Verbandes der deutschen Internetwirtschaft Eco glauben 87 % der befragten Experten, daß das Mobiltelefon in der Informationsversorgung künftig als mindestens gleichrangig mit den Printmedien und dem Fernsehen zu bewerten ist. Als wichtigste Zielgruppe für das Handy als Medium hat diese Untersuchung die Jugendlichen unter 24 Jahren ausgemacht,



Die junge Generation nutzt das Handy schon heute intensiv als Medium. Mit Sprachanwendungen wird die Bedienung komfortabler und interaktiver (Foto: O2)

aber auch Geschäftskunden spielen eine wichtige Rolle.
Aber auf welche Art und Weise wird das Handy als Medium genutzt werden? Mit der mangelhaften Benutzerfreundlichkeit von WAP scheiterte der erste Versuch, das Mobiltelefon als interaktives Massenmedium zu etablieren. SMS und die neuen MMS-Mitteilungen stoßen da schon auf eine höhere Akzeptanz.
Soll das Handy aber zum interaktiven Medium ausgebaut werden, kommen wieder die Sprachanwendungen ins Spiel. Findige Marketiers, stets auf



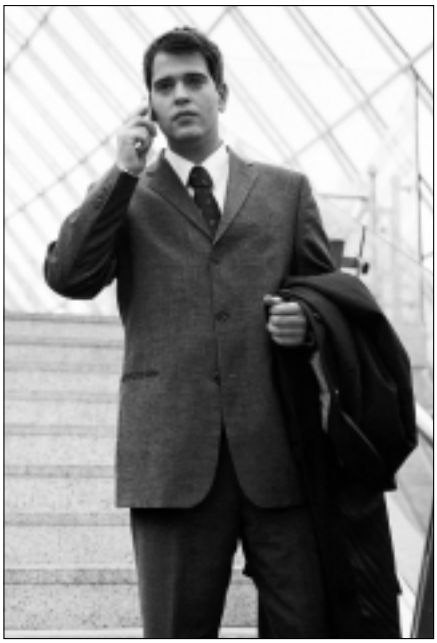
Das Voice Application Management System der VoiceObjects AG sorgt für den wirtschaftlichen Betrieb und die einfache Entwicklung von Sprachdialogsystemen sowie für die Integration der sprachverarbeitenden Techniken mit der IT-Infrastruktur

der Suche nach neuen Absatzkanälen und Plattformen für ihre Werbebotschaften, tüfteln längst an Konzepten, die die Anwendungsmöglichkeiten von Handys mit Hilfe von Sprachdialogsystemen erweitern.
Bereits heute entstehen entsprechende Voice-Portale, zunächst mit attraktiven Inhalten für die innovationsfreudige junge Handy-Generation. Die sogenannten Conversational-Marketing-Strategien zielen darauf ab, ein attraktives Werbeumfeld für handynahe Multimedia-, Entertainment- und Lifestyle-Angebote zu schaffen. Die denkbaren Geschäftsmodelle sind vielfältig und reichen von kostenpflichtigen Informations- und Unterhaltungsinhalten bis hin zu sprachmoderierten Dating- und Chat-Plattformen. Diese Services, präsentiert von einer flotten Telefonstimme oder gar der Stimme des Lieblingsstars, sollen als Einstieg der konsumfreudigen Zielgruppe schmackhaft gemacht werden.
Im nächsten Schritt, wenn sich der Voice Commerce für diese Dienste etabliert hat, werden die Marketiers das Angebot um Waren und Dienstleistungen, die wenig oder nichts mit Mobilfunk zu tun haben, erweitern, um so weitere neue Zielgruppen zu erschließen.

SMS, Klingeltöne, Handy-Logos und MMS zeigen bereits heute das Marktpotential von Mobilfunkmehrwertdiensten. Offensichtlich scheint im mobile Commerce – die bequeme bargeldlose Abrechnung über die Telefonrechnung macht's möglich – der Euro lockerer zu sitzen.
Voice Commerce, oder kurz V-Commerce, heißt nichts anderes, als mit der Sprache als Kommunikationskanal die Waren und Services noch attraktiver darzubieten. Die Präsentations-, Bestell- und Bezahlvorgänge, ja selbst die Authentifizierung (biometrisch über die Stimme) werden in einen fließenden Dialog integriert. So stört nichts die direkte, unmittelbare Ansprache auf der emotionalen Ebene, und die Kunden werden buchstäblich zum Kauf überredet.
Man braucht nicht viel Fantasie, um eine Erfolgsgeschichte wie Ebay auf Voice-Commerce-Anwendungen zu projizieren: Ließen sich online Auktionen über das Telefon nicht noch einfacher, freier, spannender, unterhaltender und mitreißender gestalten?

My Phone is my Office

Eine weitere im Internetboom propagierte Prognose bezieht sich auf die Durchdringung der Arbeitswelt durch internetbasierte Techniken. Auch das hat sich zumindest zum Teil schon bewahrheitet. E-Business-Konzepte haben die Geschäftsprozesse in einigen Branchen revolutioniert. Als Beispiel soll nur das mittlerweile weitgehend automatisierte Supply Chain Management in der Automobilindustrie dienen.
Aber auch in der täglichen Arbeitswelt sind das Internet und seine kleinen Brüder Intra- und Extranet nicht mehr wegzudenken. Viele Unternehmen haben erkannt, daß die Web-Techniken die Produktivität ihrer Mitarbeiter signifikant steigern. Unter dem Schlagwort „Mobile Business" erhalten Geschäftsreisende, Außendienstmitarbeiter und Teleworker Zugriff auf Inhalte und Anwendungen. Mit Hilfe von Sprachapplikationen könnten sie in naher Zukunft auch über das Mobiltelefon Termine und Kontaktdaten direkt verwalten oder auf mobile Office-Anwendungen und Geschäftsdaten zugreifen.



Mit Sprachdialogsystemen können Geschäftsreisende bequem per Mobiltelefon auf Anwendungen und Daten zugreifen

(Foto: O2)

Viele Marktteilnehmer – was bringt sie zusammen?

Außer den Telefongesellschaften sind an der Wertschöpfungskette für Sprachanwendungen neben den Hardwareherstellern die Softwareanbieter für die automatischen Spracherkennungs- (ASR – Automated Speech Recognition) und Sprachausgabesysteme (TTS – Text to Speech-Synthesis) sowie die Anbieter der Betriebssysteme für die Software beteiligt. Darüber hinaus spielen Systemhäuser für die Entwicklung und Implementierung der Anwendungen sowie Dienstleister für das Hosting der Anwendungen eine Rolle.
Da das Marktpotential der Sprachanwendungen zum großen Teil auf deren Bedienfreundlichkeit basiert, sind als besonders wichtige Player außerdem die VUI-Designer zu nennen, die sich um die Gestaltung anwenderfreundlicher und zielgruppengerechter Benutzerschnittstellen kümmern. Denn ohne komfortable Interaktion wird es keine Akzeptanz durch die Nutzer geben.
Aber auch die besten Designer kommen nicht ohne die Softwareanbieter aus, deren Entwicklungs- und Betriebsplattformen die technischen Grundlagen wirklich anwenderfreundlicher Anwendungen darstellen. Eine ganz zentrale Stellung in der Wertekette nehmen daher die Anbieter von Voice-Application-Management-Systemen (VAMS) ein, die die entsprechende serverbasierte Technik als Framework für die Anwendungen und zugleich die dafür benötigten Entwicklungs- und Management-Tools liefern.
VAMS integrieren die eigentlichen Sprachtechniken wie ASR und TTS mit den großteils bereits für E-Commerce- und E-Business-Applikationen aufgebauten IT-Infrastrukturen. Sie sorgen dafür, daß die Anwendungen in heterogenen Umgebungen problemlos laufen und ermöglichen die Integration ins Backend, z.B. über Schnittstellen mit Webservern, Datenbanken, CRM-, ERP-, SCM-Systemen usw. Basis für die Zusammenarbeit der Marktteilnehmer aus allen Segmenten der Wertschöpfungskette sind Standards wie VoiceXML, SALT, Datenbankmanagementstandards und Java.
VAMS verfügen über eine einfach zu handhabende Entwicklungsumgebung, die keinerlei Programmierkenntnisse erfordert. Integrierte Lifecycle- und Workflow-Managementlösungen erleichtern die Erstellung und Aktualisierung der Anwendungen.
Doch die deutliche Vereinfachung bei der Entwicklung ist nicht der einzige signifikante Vorteil beim Einsatz von Voice-Application-Management-Systemen im Carrier-Bereich: Besonders die Verwaltung und der Betrieb großer, massentauglicher Sprachanwendungen mit über tausend Ports wird durch ein serverbasiertes VAMS erst wirtschaftlich machbar. Dieses ermöglicht Administration und Monitoring ohne großen Aufwand sowie Korrekturen im laufenden Betrieb. Die TCO (Total Cost of Ownership) sinken somit beim Einsatz VAMS-basierter Systeme gerade bei großen Anwendungen im Vergleich zu herkömmlichen Umgebungen signifikant. *(bk)*