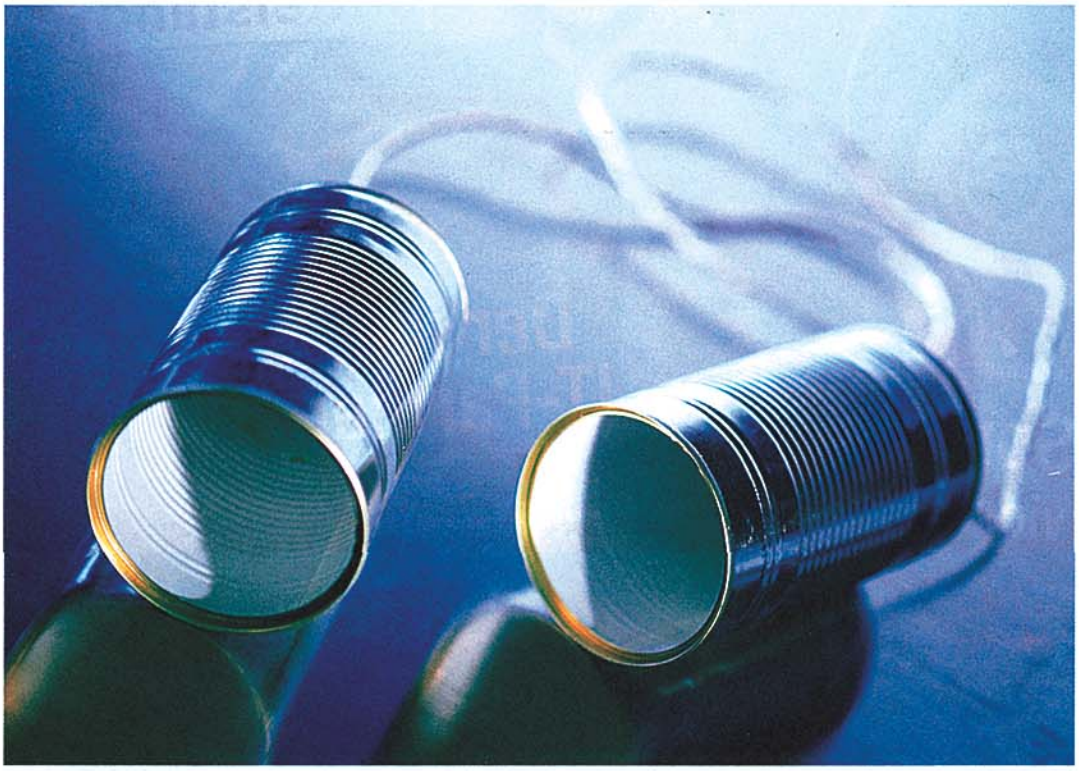
Nr. 3 vom 16. Januar 2004

Sprachanwendungen der nächsten Generation

Moderne Computer äußern Verständnis

Automatisierte Telefonapplikationen verbinden die Anrufer immer noch mittels einer „Blechstimme" aus den Anfangsjahren der Sprachsynthese. Neue Lösungen sollen dabei helfen, die Akzeptanz von Voice-Diensten zu stärken. Profitieren werden in jedem Fall die Mobil- und Festnetz-Carrier.

Von Karl-Heinz Land*



Monochrome Klangfarbe. Foto: Photo Alto

Angeklickt

Wer schon einmal mit einer Softwareanwendung im Call-Center gesprochen hat, kennt das Problem: Das gegenseitige Verständnis hält sich in Grenzen. Doch auch hier macht der technische Fortschritt nicht halt – neue Programme verstehen Sprache besser und können mit Backend-Systemen integriert werden. Kunden, Partner und Mitarbeiter lassen sich so über das Telefon an die Unternehmen binden, und Carrier versprechen sich neue Wachstumsimpulse.

SCHLECHTE ERFAHRUNGEN mit automatisierten Telefondienstleistungen haben schon viele Menschen gemacht. Die Probleme resultieren meistens aus der oft umständlichen und nicht nachvollziehbaren Benutzerführung über Tastendruck oder Spracheingabe, die den Nutzer immer nur einen kleinen Schritt weiterbringt. Hinzu kommt eine ungenügende Spracherkennung – nicht nur bei Niederbayern. Nach dem zweiten oder dritten „Ich konnte Sie leider nicht verstehen, bitte wiederholen Sie Ihre Eingabe" legt in der Regel auch der geduldigste Teilnehmer auf.

Doch diese Unzulänglichkeiten sind passé. In den vergangenen Jahren haben die Basistechnologien enorme Verbesserungen erfahren, und Standards wie Voice XML gewährleisten ein besseres Zusammenspiel der Einzelkomponenten. Zudem haben die Anbieter erkannt, dass nur eine auf die jeweilige Zielgruppe zugeschnittene, nach Usability-Gesichtspunkten gestaltete Benutzer-Schnittstelle (VUI = Voice User Interface) akzeptiert wird. Die neuen Systeme simulieren virtuelle Gesprächspartner, die auf Datenbestände im Backend der Unternehmen zurückgreifen, um Transaktionen in Echtzeit abzuwickeln und die individuellen Anfragen und Wünsche der Anrufer abbilden zu können. Dabei setzten die Sprachanwendungen auf vorhandenen E-Business-Technologien wie CRM-, Banking- oder Shop-Systemen auf.

So ist mittlerweile eine neue Generation von Sprachanwendungen herangereift: Sie zeichnet sich neben der gesteigerten Akzeptanz durch den überzeugenden Anwenderkomfort auch durch eine wesentlich höhere Wirtschaftlichkeit aus. Viele Finanzdienstleister haben bereits innovative, automatisierte Lösungen für das Telefon-Banking in Betrieb. Aber auch Versicherungen und Luftfahrtgesellschaften sehen sich nach neuen Systemen um, ebenso wie Unternehmen anderer Branchen mit einem hohen Aufkommen an Kundenanfragen, die sich standardisieren lassen.

Ein echter Sprecher im Rechner

Die Postbank etwa plant, ihr bestehendes Telefon-Banking-System in der ersten Hälfte 2004 durch moderne Sprachtechnologie zu ersetzen. Dieses Projekt zeichnet sich dadurch aus, dass durch Audio-Einspielungen mit Stimmen echter Sprecher eine bislang unerreichte Qualität erzielt wird. Zudem wird das System für ein Volumen von über 100 000 Anrufen pro Tag ausgelegt, wobei die Anwendung mehr als 700 Anrufer gleichzeitig bedienen kann. Ebenfalls bis zum Sommer will eine internationale Mobilfunkgesellschaft ein sprachgesteuertes Infotainment-Portal zunächst für die Vermarktung von Klingeltönen oder Handy-Logos anbieten. Die Voraussetzung für solche massentauglichen Lösungen sind integrierte Entwicklungs- und Management-Funktionen (so genannte Voice-Application-Management-Systeme). Diese ermöglichen – vergleichbar mit Web-Application-Servern – den wirtschaftlichen Betrieb großer Sprachanwendungen erst.

Derartige große Anwendungen, die als Sprachportale attraktive Inhalte und Services für ein Massenpublikum bereitstellen, werden die Akzeptanz und damit die Nutzung von sprachgesteuerten Telefondiensten vorantreiben. Die Marktforschungsunternehmen sind sich einig, dass dem Voice-Business ein kräftiges Wachstum bevorsteht: Datamonitor beispielsweise sagt in der Studie „Voice Portals and Applications" (Juni 2002) für den Weltmarkt der sprachverarbeitenden Lösungen ein Wachstum von heute 656 Millionen Dollar auf über drei Milliarden Dollar im Jahr 2008 voraus – ohne die damit generierten Umsätze.

Pro Tag 1,8 Millionen „Dialoge"

Die Beratungsgesellschaften Heitzig Consult und Strateco stellen in ihrer im Oktober 2003 veröffentlichten Untersuchung „Sprachapplikationen auf dem Prüfstand" fest, dass bereits heute pro Tag 1,8 Millionen Vorgänge auf deutschsprachigen Sprachdialogplattformen ablaufen. Im kommenden Jahr, so die Prognose, wird ein Viertel der deutschsprachigen Bevölkerung regelmäßig eine oder mehrere Sprachanwendungen nutzen.

Ähnlich wie beim Web werden umso mehr User das Telefon als Zugangsmedium nutzen, je mehr Content und wertvolle Services angeboten werden. Höhere Nutzerzahlen rufen wiederum neue Content- und Serviceanbieter auf den Plan. Wenn Voice-Portale dann noch untereinander verlinkt werden und die Nutzer über das Telefon auch „surfen" können, wird in nicht allzu ferner Zukunft parallel zum Internet ein „Voice-Web" entstehen. Damit könnte der Markt eine ähnliche Eigendynamik wie das Internet entwickeln und die von den Marktforschern prognostizierten Wachstumsraten deutlich übertreffen.

Umsatz rauf, Kosten runter

Gegenwärtig wird der Markt von Unternehmen angetrieben, die mittels Sprachapplikationen ihre Kosten senken, mehr Umsatz erzielen oder die Produktivität ihrer Mitarbeiter steigern wollen. Ziel ist auch hier ein schneller Return on Investment. Kostensenkungen durch effizientere Call-Center finden sich vor allem in Firmen mit vielen Kundenkontakten. E-Commerce-Anbieter und klassische Versandhändler können einen neuen Absatz- und Marketing-Kanal für bestehende Produkte und Dienstleitungen öffnen oder komplett neue Geschäftsmodelle aufstellen. Produktivitätssteigerungen versprechen sich zudem Unternehmen mit einer großen Zahl mobiler Mitarbeiter. Neudeutsch heißen die Markttreiber dementsprechend Customer-Self-Services, Voice Commerce und Mobile Workforce.

Viele Unternehmen stehen vor der Aufgabe, ihre Kunden möglichst dauerhaft an sich zu binden und gleichzeitig die Kosten für den Service zu verringern. Zu diesem Zweck wurde der Kommunikationskanal Internet in den letzten Jahren stark ausgebaut. Laut aktuellen Studien laufen aber noch immer rund 70 Prozent der Kundenkontakte über das Telefon ab

– den Call-Centern und Telefonzentralen kommt also nach wie vor eine zentrale Bedeutung zu. Diese sind jedoch wegen der hohen Personalkosten eine teure Angelegenheit. Hier findet sich ein enormes Einsparpotenzial, zumal ein Großteil der Anrufe Standardanfragen darstellen, die heute problemlos durch virtuelle Gesprächspartner bearbeitet werden können.

Die Giga Group schätzt, dass durch den Einsatz von Sprachdialogsystemen bis zu 30 Prozent der Betriebskosten eines Call-Centers eingespart werden können. Auch die Servicequalität wird verbessert: Einfache Anliegen wie beispielsweise der Reset eines Passwortes oder die Abfrage des Kontostands, aber auch komplexere Transaktionen wie Buchungen und Aktiengeschäfte erledigen automatisierte Sprachanwendungen rund um die Uhr. Dadurch werden die „echten" Call-Center-Agenten entlastet, so dass sie sich um individuelle Kundenwünsche kümmern und beratungsintensive (und umsatzsteigernde) Gespräche führen können.

Mobilfunk-Mehrwertdienste wie Handy-Klingeltöne, -Spiele oder -Logos haben ihr Marktpotenzial im Mobile Commerce bereits bewiesen. Doch die bis dato umständliche Bedienung über SMS und WAP muss nicht sein – eine komfortable und zielgruppengerechte Kunden-„Ansprache" im Mobile-Commerce erleichtert die Handhabung der Auswahl-, Bestell- und Bezahlvorgänge. Zur Zielgruppe passende, sogar personalisierbare Stimmen in einer attraktiven, zum Beispiel mit den neuesten Chart-Hits hinterlegten akustischen Umgebung halten die Anrufer in der Anwendung und überreden sie zum Kauf.

Laut aktuellen Studien laufen noch immer rund 70 Prozent der Kundenkontakte über das Telefon ab – den Call-Centern und Telefonzentralen kommt also nach wie vor eine zentrale Bedeutung zu.

Voice-Portale öffnen sich

Erste Voice-Portale mit Multimedia-, Infotainment- und Lifestyle-Angeboten für die jugendliche Handy-Generation entstehen bereits und schaffen nicht zuletzt auch ein attraktives Werbeumfeld. Sind sie erst einmal als Massenmedium akzeptiert, wollen die Anbieter darüber hinaus auch „Handy-ferne" Waren und Dienstleistungen – von der Kinokarte bis zur Urlaubsreise – anbieten.

Parallel dazu werden weitere Zielgruppen erschlossen, etwa mit Geschäftskundenportalen, die aktuelle Informationsinhalte personalisiert bereitstellen.

Um auch außerhalb des Büros produktiver arbeiten zu können, sollen die Mitarbeiter überall und jederzeit auf E-Business-Anwendungen und Daten zugreifen können. Die Sprachsteuerung bietet auch hier einige Vorteile, denn zur verbalen Interaktion genügt ein simples Telefon oder Handy. Eine Datenverbindung muss nicht hergestellt werden, falls das Notebook oder der PDA mal im Büro geblieben sind. Als Beispiel soll der Vertriebsmitarbeiter einer Versicherung dienen: Mit einem Sprachdialogsystem erhält er über sein Handy Zugang auf die diversen branchenspezifischen Anwendungen. So kann er kurz vor einem Kundentermin noch im Auto die nötigen Daten in das System eingeben. Dieses teilt ihm daraufhin die Kundenhistorie oder individuell berechnete Tarifmodelle mit. Zusätzlich lassen sich Termine und Kontaktdaten verwalten sowie Geschäftsdaten abfragen und verschiedene Anwendungen wie E-Mail aufrufen.

Egal, welches Unternehmen die sprachbasierenden Services für Kunden, Partner oder Mitarbeiter auch anbietet, eine Branche ist immer der Gewinner: die Telekommunikationsgesellschaften. Weil jedesmal ein Telefongespräch abläuft, profitieren von dem Trend sowohl die Festnetz- als auch Mobilfunkbetreiber. Die Carrier erwarten, dass ihre stagnierenden oder gar sinkenden Einnahmen in ihrem Kerngeschäft, der Telefonie, mit dem Durchbruch von Sprachanwendungen wieder steigen. Zudem können sie lukrative, kostenpflichtige Mehrwertdienste anbieten und im Voice Commerce mitmischen – und damit den durchschnittlichen Umsatz pro Kunde (ARPU = Average Revenue Per User) als entscheidenden Wettbewerbsfaktor steigern.

Und obwohl es auf den ersten Blick widersprüchlich scheint: Auch die Mobilfunkbranche hofft, mit den Sprachanwendungen endlich einen Motor für ihre mobilen Datendienste gefunden zu haben. In einer Marktanalyse bezeichnet Forrester Research Sprachanwendungen als „Trojanisches Pferd", mit dem die Mobilfunkbetreiber dem Kunden die gewinnträchtigen Datendienstleistungen sozusagen auf dem Umweg über die Sprache schmackhaft machen können. Indem die Kunden Applikationen bequem über den Modus „Sprache" und das Medium Telefon steuern, werden sie schrittweise an die multimedialen Endgeräte der nächsten Generation herangeführt.

Bedienkomfort entscheidet

Künftig lassen sich verschiedene Interaktionen integrieren, und der Benutzer wird selbst entscheiden können, ob er lieber Sprache, Tastaturen, Displays, Touchscreens oder eine Stylus-Handschriftenerkennung nutzen möchte. Multimodalität als Voraussetzung für Multimedia könnte sich nicht zuletzt für die TK-Anbieter als überlebenswichtig erweisen: Egal, welche Services in Zukunft angeboten werden, entscheidend ist der Bedienkomfort immer komplexer werdender Anwendungen auf schrumpfenden Endgeräten. Ist das nicht gegeben, besteht die Gefahr, dass Kunden moderne UMTS-Dienste links liegen lassen und die Anbieter auf ihren hohen Investitionen sitzen bleiben. *(ajf)* ←

*Karl-Heinz Land ist Vorstandsvorsitzender der Voiceobjects AG in Bergisch Gladbach.