

Betriebswirtschaftliche Evaluation von Spracherkennungssystemen

Diplomarbeit

zur Erlangung des Grades eines Diplom-Ökonomen der
Wirtschaftswissenschaftlichen Fakultät der Universität Hannover

vorgelegt von

Carsten Köppe



Erstprüfer: Prof. Dr. Michael H. Breitner

Hannover, den 20. Oktober 2003

Inhaltsverzeichnis

Abbildungsverzeichnis	4
Verzeichnis benutzter Abkürzungen.....	5
1 Einleitung	6
2 Technische Grundlagen von Spracherkennungssystemen.....	8
2.1 Anforderungen an die automatische Spracherkennung.....	8
2.1.1 Probleme bei der Spracherkennung.....	8
2.1.2 Leistungskriterien von Spracherkennungssystemen.....	9
2.2 Vorverarbeitung	11
2.2.1 Analog-Digital-Umsetzung	12
2.2.2 Signalanalyse	12
2.3 Spracherkennung.....	15
2.3.1 Verfahren zur Spracherkennung	15
2.3.2 Akustische Modellierung	16
2.3.3 Linguistische Modellierung.....	17
2.4 Entwicklung offener Standards: VoiceXML und SALT.....	18
2.5 Blick nach vorn	20
3 Einsatzgebiete der automatischen Spracherkennung.....	21
3.1 Sprachsteuerung.....	22
3.1.1 Maschinensteuerung und Datenerfassung	22
3.1.2 Steuerung verschiedenster Geräte und Vorgänge	24
3.1.3 Sprachsteuerung im Auto.....	25
3.2 Diktiersysteme	25
3.2.1 Produkte für den Massenmarkt.....	25
3.2.2 Spezielle Lösungen.....	26
3.3 Telephonieapplikationen.....	27
3.3.1 Steuerung des Telefons	27
3.3.2 Call Center Automatisierung	27
3.3.3 Voice Portale	28
3.4 Zugangssysteme und Sicherheitstechnik	30
3.5 Fremdsprachenbereich	31
3.5.1 Übersetzung gesprochener Sprache	31
3.5.2 Erlernen von Fremdsprachen	31
3.6 Behindertenunterstützung und therapeutischer Bereich.....	32

4	Kosten, Nutzen und Potentiale	33
4.1	Prinzip der Kosten-Nutzen-Analyse	33
4.2	Erleichterung von Arbeitsabläufen.....	34
4.3	Mehr Komfort vers. Sicherheit im Auto.....	35
4.4	Wie sieht das Haus der Zukunft aus?.....	36
4.5	Diktiersysteme für Privatanwender und Speziallösungen.....	38
4.6	Chancen durch Voice Portale	40
4.6.1	Verbesserung der Kundenzufriedenheit.....	40
4.6.2	Neue Möglichkeiten durch Voice Portale.....	41
4.6.3	Anforderungen an erfolgreiche Voice Portale.....	42
4.6.4	Kostenreduzierung in Call Centern durch Sprachsysteme.....	44
4.7	Vorteile biometrischer Zugangssysteme	45
4.8	Neue Kommunikationsmöglichkeiten durch Übersetzungstools	46
4.9	Erleichterungen für Behinderte	46
5	Der Markt für Spracherkennungssysteme.....	47
5.1	Anbieter und Geschäftsmodelle	47
5.1.1	Definition Geschäftsmodell	47
5.1.2	Softwareentwickler	48
5.1.3	Betreiber von Voice Portalen.....	50
5.2	Konzentrationsprozesse	52
5.3	Zukunftsaussichten.....	53
6	Fazit und Ausblick.....	56
	Literaturverzeichnis	58
	Anhang	62

Kapitel 1

Einleitung

Der Supercomputer HAL 9000, der in Stanley Kubricks oscarprämiierten Film „2001: Odyssee im Weltraum“ das Kommando eines Raumschiffs übernimmt, zeigte den Zuschauern die Vision einer intelligenten, sprechenden Maschine.¹ Nun leben wir im Jahr 2003 und können überprüfen, was aus der Vision Kubricks geworden ist. Einen sprechenden Supercomputer wie den HAL 9000 haben wir noch nicht, aber es gibt verschiedene Anwendungen von Sprachtechnologie, die viel versprechende Schritte in diese Richtung einläuten.

Im Rahmen dieser Arbeit sollen die unterschiedlichen Spracherkennungssysteme (SES) herausgearbeitet werden. Dabei werden sie weniger auf ihre technischen Eigenschaften untersucht, vielmehr stehen betriebswirtschaftliche Überlegungen im Vordergrund. Es geht also im Kern der Arbeit um die Frage nach dem Nutzen von SES. Haben die Systeme einen Nutzen für potentielle Anwender? Oder sind sie nur technische Spielerei? Und lässt der zusätzliche Nutzen die entstehenden Kosten rechtfertigen?

Mögliche Potentiale von SES sollen ebenso untersucht werden wie Hindernisgründe, die sowohl technischer als auch rechtlicher Art sein können. Und wie sieht es mit der Akzeptanz durch mögliche Kunden aus? Demnach sind auch die Eigenschaften und individuellen Lösungsansätze von SES, welche die Anwenderfreundlichkeit bestimmen, zu durchleuchten.

In den vergangenen Jahrzehnten gab es immer wieder Zeiten voller Euphorie, in denen der Durchbruch in der Sprachtechnologie für die nahe Zukunft vorher gesagt wurde. Diese wurden jedoch schnell von Zeiten des Pessimismus abgelöst, wenn sich zeigte, dass die technischen Möglichkeiten noch nicht so weit waren, dass komfortable Anwendungen angeboten werden konnten. Für Systeme mit hohen Fehlerraten hatten die Konsumenten kein Verständnis. Doch wie sieht es heute aus? Stehen wir heute vor einem Durchbruch? Oder wird dieser der Sprachtechnologie nie gelingen?

Im Rahmen dieser Abhandlung werden daher realisierte Anwendungen von Spracherkennung vorgestellt, die hohe Erkennungsraten aufweisen, und es werden

¹ mehr zum Film unter <http://www.cyberkino.de/entertainment/kino/101/101850.html>

Dialogsysteme gezeigt, die einem zwischen menschlichen Dialog erstaunlich nahe kommen.

Um die Sprachtechnologie weiter voranzutreiben und vielleicht schon bald zum Durchbruch zu verhelfen, bedarf es Unternehmen, die bereit sind, in die Entwicklung dieser Technologien zu investieren. Daher soll der Markt für Spracherkennungsprodukte näher betrachtet. Es werden verschiedene Hersteller mit unterschiedlichen Geschäftsmodellen ebenso wie vollzogene Konzentrationsprozesse betrachtet.

Die zur Fragestellung dieser Arbeit bisher vorhandene Literatur beschränkt sich hauptsächlich auf eine Abhandlung von Susen [1999] und einen Aufsatz von Gallwitz u. a. [1999], die beide im Jahr 1999 Einsatzmöglichkeiten und Nutzen von Spracherkennungssystemen untersucht haben. Aufgrund der fortschreitenden technischen Entwicklung sind einige ihrer Ergebnisse jedoch nicht mehr zutreffend und es macht Sinn, auf Grundlage des heutigen Standes Einsatzmöglichkeiten, Kosten und Nutzen der automatischen Spracherkennung erneut zu untersuchen. Es zeigt sich, dass einige der damaligen Visionen schon umgesetzt wurden, dass andere Einsatzmöglichkeiten jedoch nicht weiterverfolgt wurden.

Zur betriebswirtschaftlichen Evaluation von Spracherkennung gehört auch die Analyse der Zukunftsaussichten, der sich das letzte Kapitel dieser Abhandlung widmen wird.

Zuvor werden jedoch die technischen Grundlagen der automatischen Spracherkennung kurz vorgestellt. Außerdem soll verdeutlicht werden, wo die Schwierigkeiten bei der Erkennung gesprochener Sprache sind, und worin diese begründet sind.

Kapitel 6

Fazit und Ausblick

Die vorangegangenen Kapitel haben deutlich gezeigt, welche Einsatzmöglichkeiten Spracherkennungssysteme bieten. Die Anwendungsfelder sind vielfältig und reichen von Sprachsteuerung und Diktiersoftware über Call Center Automatisierung und Voice Portale bis hin zu Sicherheitstechnik und Übersetzungstools. SES können dem Anwender demnach in vielen Bereichen des betrieblichen Schaffens, aber auch im privaten Bereich von hohem Nutzen sein.

Auch wenn einige Anwendungen als technische Spielerei gewertet werden können, steckt in anderen dennoch großes Potential. Eine Steuerung verschiedenster Geräte per Spracheingabe wird in Zukunft vermutlich ebenso alltäglich sein, wie die Sprache mit automatisierten Call Centern und Voice Portalen. Gerade im Bereich der Voice Portale steckt beträchtliches Potential, da sie dem Nutzer ermöglichen, unabhängig von PC und klassischen Internetzugängen Informationen aus dem World Wide Web abzurufen und Transaktionen durchzuführen. Mit intelligenten Voice Portalen könnte das Telefon zum neuen Zugang zum Internet werden und auf diese Art und Weise neue Zielgruppen für das Internet gewinnen.

Die Call Center Automatisierung bietet Unternehmen die Gelegenheit, Kosten einzusparen und gleichzeitig die Kundenzufriedenheit zu erhöhen, indem das Unternehmen 24 Stunden am Tag für die Fragen und Wünsche der Kunden offen ist, lästige Warteschleifen entfallen und eine Vielzahl von Informationen per Telefon abrufbar sind. Es sollten aber auch die durch SES möglichen Verbesserungen für Behinderte nicht unerwähnt bleiben. Auch wenn es sich hierbei um einen sehr kleinen Markt handelt, ist der Nutzen für den einzelnen umso größer.

Die beschriebenen Einsatzmöglichkeiten und Potentiale und der zu erwartende Nutzen für die Anwender zeigen die enormen Chancen für die Hersteller von Systemen zur automatischen Spracherkennung. Auch wenn ihre Entwicklungen bisher kaum Gewinne abgeworfen haben, kann davon ausgegangen werden, dass der Durchbruch unmittelbar bevorsteht. Einige der getesteten Anwendungen erwecken große Zuversicht.

Aufgabe der Entwickler bleibt es aber, stets im Auge zu behalten, welche Art von Anwendungen dem Anwender einen zusätzlichen Nutzen bringen, und wie diese Anwendungen beschaffen sein müssen, damit sie akzeptiert werden.

Im Segment der Voice Portale sind bisher zwar einige Anfänge gemacht, es gibt hier aber noch große unausgeschöpfte Möglichkeiten, erfolgreiche Voice Portale zu schaffen, die den Nutzern Vorteile verschaffen und Freude bringen. Folglich bestehen auch für Neueinsteiger Chancen, sich eine bedeutende Rolle unter den Voice Portalen zu erarbeiten.

Der Markt für Sprachtechnologie ist ein stark wachsender Markt, der nicht nur den Herstellern und den Betreibern von Call Centern und Voice Portalen oder den Anwendern der Systeme Vorteile verschaffen kann. Auch Dienstleister wie Telekommunikationsunternehmen können an diesem Wachstum partizipieren, da Voice Portale zusätzliche Gesprächsminuten und damit Umsätze generieren, ebenso wie Service-Provider und Hardwarehersteller.

Inwieweit die Visionen des sprechenden Computers und vom Haus der Zukunft bald Wirklichkeit werden, lässt sich noch nicht mit Bestimmtheit voraussagen. Zum einen muss die Industrie solche Entwicklungen mutig vorantreiben, zum anderen müssen neue Technologien vom Konsumenten akzeptiert werden. Es bleibt abzuwarten, was geschieht.

Aber vielleicht werden wir uns ja doch irgendwann einmal ganz selbstverständlich mit unserem Computer unterhalten, wie es Captain Jean-Luc Picard auf der Enterprise macht, wenn er seinen „Tee, Earl Gray, heiß“ beim Nahrungs-Replikator bestellt.