

1. Einführung

1.1. Das Unternehmen

Die Firma Ebert-Ingenieure ist in der Baubranche mit deutschlandweit 300 Mitarbeitern tätig.

Teilweise werden weltweit Bauprojekte abgewickelt. Dies umfasst, Planung, Konzeption, Durchführung und Abrechnung von Bauprojekten.

1.2. Das Projekt

In dieser Branche entsteht sehr viel Schreibaufwand. Für die Zukunft soll an schreibaufwendigen Arbeitsplätzen, eine Spracherkennungssoftware installiert werden, die Sprache in Text umsetzt.

Anwendungsgebiete sind z. B. die Rechtsabteilung, Vertragsgestaltung, oder die Protokollierung von Konferenzen (Erfassung auf ein digitales Diktiergerät zur späteren Verarbeitung)

Für diese Aufgaben werden Anwendungsprogramme wie:

- OfficeWare (Office Management Software, dient zum internen Dokumentenaustausch)
- Microsoft Word (für Standard-Briefe etc.)
- Texteditor (für schnelles editieren von Ideen ohne Formatierung)
- Outlook (für externe Kommunikation) eingesetzt.

2. Ist-Aufnahme

2.1. Vorhandene Hardware

Für den Einsatz der Software kommen 80 Workstations in Frage.

Folgende Hardwareanforderungen sind für die Software Voraussetzung:

Mindestens,

- Intel PII 400MHz oder vergleichbares
- 128MB RAM
- Windows kompatible 16Bit- Soundkarte
- 500 MB freier Festplattenspeicher vorhanden

Mit Hilfe einer Liste, die ich vom Administrator erhalten habe, prüfte ich welche Arbeitsstationen den Anforderungen entsprachen.

Ergebnis meiner Recherche:

	Ja	Nein
Intel Pentium \geq 400 MHz	62	18
Arbeitsspeicher 128 MB – 256 MB	49	31
Freier Festplattenspeicher mind. 500 MB	80	0
Windows- kompatible 16- Bit Soundkarte	20	60
USB Schnittstelle	80	0
USB Headset	0	80

Durch diese Aufstellung konnte ich mir einen besseren Überblick über die aktuell vorhandene Hardware im Unternehmen verschaffen.

2.2. Vorhandene Software

Auf allen Workstations sind Office Anwendungen, wie

- Word
- Excel
- MS- Outlook
- OfficeWare (Dokument-Management-System)

sowie diverse branchenspezifische Softwarelösung auf einer Windows-2000-Plattform installiert.

2.3. Netzwerk

Da ein internes Netzwerk (Intranet) vorhanden ist, kann in einem späteren Schritt eine Installation der Sprachsoftware auch im Netzwerk realisiert werden.

Das hat den Vorteil, dass hier zentral das Profil eines jeden Benutzers gespeichert werden kann.

Somit besteht die Möglichkeit, sich an einer anderen Workstation anzumelden und das jeweilige Benutzer- Sprachprofil zu laden.

Dies führt zu mehr Flexibilität und Mobilität.

Die Netzwerkverwaltung erfolgt zum größten Teil von Nürnberg aus.

3. Soll-Zustand

3.1. Anforderungsprofil

In der Unternehmensleitung wurde beschlossen, Lizenzen für ein sprachgesteuertes Programm zu erwerben das die Arbeit an den Workstations wesentlich erleichtern soll.

Hauptziel des Unternehmens ist es, bei allen unter Punkt 2.2 genannten Anwendungen eine sprachgesteuerte Texterkennung zu realisieren.

Dies bringt Zeitersparnis und somit Kostensenkung.

Weitere Anforderungen an das Programm sind:

- Das gesamte System, auch die branchenspezifische Software, soll mit einer Spracherkennungssoftware gesteuert werden.
- Die Installation und Benutzung der Spracherkennungssoftware sollte unkompliziert sein.

Das bedeutet, neue Benutzer sollten dieses Programm einfach anwenden können, ohne langwierige Einarbeitung und große Programmkonfigurationen durchführen zu müssen.

Die Eingabe in Word oder Texteditor etc. sollte möglichst komfortabel und ohne größere Probleme vonstatten gehen.

3.2. Budgethöhe

Als Budgethöhe für den Test wurde ein Betrag von € 1.600,- festgelegt.

Dieser Betrag soll den Stundensatz eines Mitarbeiters, die Beschaffung der Testsoftware und die Nutzung des PCs für den Testzeitraum decken.

4. Produkt-Recherche

4.1. Internet-Recherche

Die Suche nach den geeigneten Programmen erwies sich schwieriger als gedacht, denn die Produkte sind nur auf wenige Anbieter beschränkt.

Nach längerer Suche stellte sich heraus, dass aktuell nur drei Programme auf dem freien Markt angeboten werden.

Diese Programme sind:

- **IBM ViaVoice Pro 10**
- **linguatec VoicePro 10 USB-Edition** (enthält USB- Headset)
- **Dragon NaturallySpeaking Professional XP Edition**

Zur **linguatec VoicePro 10 USB-Edition** Software ist zu sagen, dass diese auf Basis der **IBM**-Programmierung aufgebaut ist und nur noch kleine herstellerspezifische Einbindungen hat.

Deshalb wurde diese Software bei meinen Tests nicht berücksichtigt.

4.2. Fachzeitschriften-Recherche

Im Projektzeitraum erschienen aktuelle Tests in bekannten Fachzeitschriften. Die Meinungen und Testergebnisse zog ich für einen späteren Vergleich heran.

Dabei kam es mir hauptsächlich auf die Erkennungsgenauigkeit und Verarbeitungsgeschwindigkeit der Programme an.

4.3. Telefon-Recherche

Bei meiner Recherche im Internet stieß ich auf einen Vertrieb, der die beiden Softwareprogramme verkauft.

Ich setzte mich mit dem Unternehmen telefonisch in Verbindung.

Es stellte sich heraus, dass der Verkaufsleiter schon seit längerem mit einem Spracherkennungsprogramm arbeitet.

Meine Fragen stellte ich hinsichtlich der Bedienbarkeit, der Effektivität und der Kosten.

Mir wurde unter anderem erklärt, dass es Zeit braucht um ein solches Programm zu bedienen bzw. kennen zu lernen.

Einen Favoriten hatte der Verkaufsleiter für sich schon gefunden. Über das Gespräch machte ich mir Notizen, um sie später ebenfalls auszuwerten und mit meinen Ergebnissen zu vergleichen.

Für **Dragon NaturallySpeaking** und **IBM ViaVoice 10** konnte der Verkaufsleiter mir leider keine freien Testversionen zur Verfügung stellen, aber gegen einen kleinen Testpreis wäre dies möglich.

4.4. Produktanforderungen

4.4.1. Hardware

Die Hardwareanforderungen der Spracherkennungsprogramme waren, gemessen am aktuellen Leistungsstandard, nicht zu hoch.

Die Hersteller gaben an:

- Mindestens einen Pentium II mit 400 MHz
- 128 MB RAM
- 500 MB freien Festplattenspeicher
- 16Bit- Soundkarte, Windows kompatibel
- Headset oder Stand-alone Mikrofon
- Evtl. Lautsprecher zur akustischen Wiedergabe
- CD-ROM-Laufwerk.

Mein Testrechner erfüllte diese Kriterien.

Zum Einsatz kam ein digitales USB- Headset (Soundkarte integriert, Plug & Play fähig), da der Test unter optimalen Bedingungen durchgeführt werden sollte.

4.4.2. Betriebssystem/Software

Zur Verwendung der Spracherkennungssoftware können folgende Betriebssysteme bzw. Versionen eingesetzt werden:

- Windows 98 SE
- Windows ME
- Windows 2000
- Windows XP (Home/Pro Edition)

Zur Texteingabe kann der im Programm enthaltene Texteditor, oder auch Office-Programme wie Microsoft Word oder StarOffice verwendet werden.

4.4.3. Merkmale

Die Programmunterschiede sind nicht sehr groß, sie bieten fast das gleiche Leistungsspektrum.

Um hier diese Dokumentation nicht mit Leistungsmerkmalen zu überhäufen, habe ich mich nur auf die wesentlichen Punkte beschränkt.

Bei den unten aufgelisteten Programmen handelt es sich um Einzelplatzversionen. Eine Netzwerklösung bietet nur **Linguatec VoicePro 10**, da Sie aber nicht zum Testeinsatz kam, gehe ich nicht weiter darauf ein. Begründung ist unter Punkt 4.1 nachzulesen.

	IBM ViaVoice Pro10	Dragon NaturallySpeaking
Steuern von Anwendungen	X	X
Personalisieren von Vokabularen mit Namen und Fachausdrücken durch Dokumentanalyse	X	X
Texterstellung in versch. Anwendungen, MS-Office, Texteditor, StarOffice,	X	X
Wortschatzgröße (inkl. Vokabulärerweiterung)	1.000.000	810.000
Diktieren, formatieren, editieren, korrigieren	X	X
Browser, Email und Betriebssystem steuern	X	X
Transkription (Digitales aufnehmen auf Diktiergerät) und als Text einspielen.	X	X
Trainingsdauer Spracherkennung) ca. in Stunden (Subjektive Meinung) - Hochrechnung	25	5
Kosten	179,00 €	926,00 €

5. Anfordern der Testsoftware

5.1. Telefonische Anfrage

Vor Bestellung der Software klärte ich telefonisch die Vertragsbedingungen mit dem Lieferanten.

Zur gegenseitigen Absicherung wurde folgendes festgelegt:

- Rückgaberecht innerhalb 14 Testtagen
- Fälligkeit des Verkaufspreises nach 14 Testtagen
-€ als Leih-, Sicherungsgebühr für 14 Testtage
- Versandkosten trägt der Empfänger

Im Anschluss legte ich die Vertragsbedingungen dem Abteilungsleiter vor, der daraufhin per Fax die Bestellung mit den oben genannten Vereinbarungen abschickte.

6. Installation

6.1. Treiber-Deinstallation und Installation

Um Treiberkonflikte mit der Software des digitalen USB-Headsets zu vermeiden entfernte ich die Soundkarte und die dazugehörigen Treiber aus dem Testrechner.

Die Empfehlung des Herstellers für das Headset war, die Software inkl. der Treiber **vor** dem anschließen des Headsets zu installieren.

6.2. IBM-Software

Die Installation verlief reibungslos, die anschließende Konfiguration erwies sich jedoch als undurchsichtig und langwierig.

6.3. Dragon-Software

Die Installation erfolgte unter gleichen Voraussetzungen.

Schon bei der Installation rechtfertigte das Programm seinen hohen Preis.

Sie verlief reibungslos und die Konfigurationseinstellungen waren klar und deutlich programmiert.

7. Konfiguration der Software

7.1. IBM-Software

Die Sprachkonfiguration war sehr aufwendig, da die Erkennung der gesprochenen Wörter teilweise überhaupt nicht funktionierte.

Ebenso fehlte ein Einführungs-Guide um den Anwender leichter ins Programm einzuführen.

Der Zeitaufwand für die Konfiguration lag bei ca. 1 Stunde, um das Programm für seine eigentlichen Aufgaben vorzubereiten. Die Erkennungsgenauigkeit bei der Konfiguration lag hier bei ca. 50%.

7.2. Dragon-Software

Die Konfiguration war klar und verständlich programmiert, und die Einführung in das Programm wurde mit einem Konfigurations-Guide fortgeführt.

Zuerst wurden vorprogrammierte Beispiele gezeigt, anschließend konnte man diese Beispiele gleich in der Praxis umsetzen.

Meiner Einschätzung nach lag die Erkennungsgenauigkeit bei der Konfiguration von Anfang an bei ca. 90%.

8. Softwaretest

8.1. IBM-Software

Das Erstellen von Textdokumenten mittels Sprache gestaltete sich genauso schwierig wie die Konfiguration.

Trotz des höheren Vokabularumfangs wurden sehr viele Wörter nur schlecht, oder teilweise überhaupt nicht erkannt.

Nicht erkannte Wörter mussten manuell trainiert werden, was sehr zeitraubend war.

Meiner Einschätzung nach lag hier die Erkennungsgenauigkeit beim Diktieren ebenso bei ca. 50%.

Der Softwaretest dauerte ca. 2,5 Stunden.

Um das Programm einigermaßen vernünftig anwenden zu können, werden zirka 25 Stunden benötigt (Hochrechnung).

8.2. Dragon-Software

Das Erstellen von Texten mittels Sprache war fast ein " Kinderspiel ".

Zirka 90% der eingegebenen Wörter wurden sofort erkannt, denn diese Wörter waren alle schon im Vokabular enthalten und auch besser programmiert. Somit entfiel die aufwändige manuelle Nachbearbeitung (Sprachtraining).

Waren jedoch einzelne Wörter nicht im Vokabular enthalten, konnten diese manuell eingegeben und mit einem leicht zu bedienenden Aufnahmerekorder aufgenommen werden.

Der Softwaretest dauerte ca. 1 Stunde. Danach lag meiner Einschätzung nach hier die Erkennungsgenauigkeit beim Diktieren ebenso bei ca. 90%.

9. Nutzwertanalyse

9.1. Nutzen

Anhand der unten aufgeführten Tabelle ist zu erkennen, wie die Bewertungen im Einzelnen vergeben wurden.

Dabei wurde der Schwerpunkt auf das Diktieren, Formatieren und Korrigieren gesetzt. Den zweitgrößten Anteil besitzt die Trainingsdauer der Programme, denn was nützt ein Programm, das sozusagen erst lange auf die Stimme „programmiert“ werden muss.

Hierbei schnellen dann die Kosten in die Höhe.

Das Leistungsspektrum beider Programme hält sich fast die Waage, wobei sie sehr stark in der Trainingszeit differieren.

Ein Unterschied ist auch in der Erkennungsleistung zu bemerken, obwohl **IBM ViaVoice** mehr Vokabeln zur Verfügung stellt.

Stand 15.02.03	Gewichtung (Firma)↓		IBM ViaVoice Pro10	Dragon NaturallySpeaking	
Steuern von Anwendungen	8	2	16	3	24
Diktieren, formatieren, editieren, korrigieren	65	2	130	4	260
Browser, Email und Betriebssystem steuern	5	2	10	3	15
Transkription (Digitales aufnehmen auf Diktier- gerät) und als Text einspielen.	2	2	4	3	6
Trainingsdauer (Spracherkennung) ca. in Stunden Hochrechnung (Subjektive Meinung)	25	2	50	4	100
Bewertungszahlen	100		200		405

1 = Nicht | 2 = Schlecht | 3 = Mittel | 4 = Gut

1 = Nicht | 2 = Schlecht | 3 = Mittel | 4 = Gut

10. Kosten

10.1. Aufstellung/Vergleich

IBM ViaVoice Pro 10	25 Std. x 40 €/h	1000,--€
Lizenz		+ 179,--€
		<u>1179,--€</u>
Dragon NaturallySpeaking	5 Std. x 40 €/h	200,--€
Lizenz		+ 926,--€
		<u>1126,--€</u>

Kosten**Ebert-Ingenieure Nürnberg**

Projekt: EDV 2003

Spracherkennungssoftware – Kostenaufstellung**Berechnungsvoraussetzungen**für alle MA außer CAD
Stundensatz intern80 MA - PCs betroffen
40,00 € Std.**Hardwarevoraussetzungen**Prozessor
Arbeitsspeicher
Soundkarte
USB- AnschlussPII 400MHz muss sein
256 MB sollte sein
keine
muss sein

	PC	Einzelbetrag	Summe
Anschaffung neuer Workstations			
Anzahl PCs unter PII 400	18	600,00 €	10.800,00 €
Anschaffung USB- Headset (integr. Soundkarte)	80	70,00 €	5.600,00 €
Workstation-Kosten			16.400,00 €

Aufrüstung des Arbeitsspeichers

Anzahl PCs ohne ausreichenden Speicher	31	25,00 €	775,00 €
Einbau u. Konfigurationsaufwand pro Speicher 0,25 Std.	31	40,00 €	310,00 €

Arbeitsspeicher-Kosten			1.085,00 €
-------------------------------	--	--	-------------------

Softwareinstallation, Konfiguration, Schulung

Installations- Konfigurationsaufwand pro PC 0,75Std.	80	40,00 €	2.400,00 €
Schulungsaufwand pro MA 1Std.	80	40,00 €	3.200,00 €

Softwareinstallations-, Konfigurations-, Schulungskosten			5.600,00 €
---	--	--	-------------------

Hardwareinvestition/Konfigurations-, Schulungskosten Summe			23.085,00 €
---	--	--	--------------------

Softwarekosten-Konfigurationskosten		80 Lizenzen		
		IBM ViaVoice		Dragon NaturallySpeaking
Lizenzkosten	179,00 €	14.320,00 €	615,23 €	49.218,40 €
Sprachtrainingsaufwand pro Mitarbeiter 25 Std.		80.000,00 €	5 Std.	16.000,00 €
Hardwareinvestition/Konfigurations-, Schulungskosten		23.085,00 €		23.085,00 €
Gesamtkosten		117.405,00 €		88.303,40 €

Preisstaffelung Lizenzen		
IBM ViaVoice		Dragon NaturallySpeaking
USER	netto	netto
1 bis 4	179,00 €	926,00 €
5 bis 50	179,00 €	639,20 €
51 bis 100	179,00 €	615,23 €
101 bis 250	179,00 €	599,25 €

Die Software **IBM ViaVoice Pro 10** ist zwar auf den ersten Blick kostengünstiger, aber der entscheidende „Kostensprung“ entsteht beim Trainieren des Programms.

Die Software **Dragon NaturallySpeaking** dagegen schlägt mit 926,00 € zu Buche, dies ist aber mit der erheblich besseren Erkennungsgenauigkeit und der viel weniger aufzuwendenden Trainings/Übungszeit zu kompensieren.

Das Programm **Dragon NaturallySpeaking** scheint auf den ersten Blick teurer.

Bei genauerer Analyse und Verwendung kann es durch extrem gute Erkennungsgenauigkeit und kurze Trainingszeit mehr überzeugen als das Konkurrenzprodukt.

11. Fazit

Vorab, "ein leistungsstarkes Auto mit allen Extras hat auch seinen Preis";

Dennoch ist es wahrscheinlich, dass es günstiger ist als ein leistungsschwacher Oldtimer mit einem sehr hohen Benzinverbrauch.

Das, was auf den ersten Blick teuer erscheint, muss nicht gleich das Teuerste sein. Im Gegenteil, nach Testende erwies sich **Dragon NaturallySpeaking** sogar als die günstigere und bessere Software.

Um eine hohe Erkennungsgenauigkeit zu erreichen, sollten die Programme bzw. die Wörter öfters trainiert werden.

Bei einem Programm mehr, einem anderen Programm weniger.

Ein geübter zehn Fingerschreiber ist zuerst auf herkömmliche Weise besser bedient, als mit einer untrainierten Spracherkennungssoftware.

Aber bringt dieser Zeit, Training und etwas Geduld mit, kann sogar er genauso schnell die Texte per Sprachsoftware eingeben.

Für Ungeübte, ist die Spracherkennungssoftware das optimale Werkzeug zur Erstellung eines Textes.

Sprachsoftwareprogramme werden von Version zur Version immer leistungsstärker, genauer und effektiver.

Zukunftsvisionen wie die SCIFI-Serie "Star Trek" wo man mit dem Computer "spricht", werden in naher Zukunft vielleicht schon Wirklichkeit.