

Erste Ergebnisse der Analyse juristischer Texte für Spracherkennungssysteme

Jörg H. Blankenburg und
Manfred Weihermüller †

Motivation

Auf den letzten drei EDV-Gerichtstagen in Saarbrücken wurde bereits deutlich, daß das Interesse der Juristen an einer anderen Art der Interaktion mit dem Computer als mittels der Tastatur, nämlich der Spracheingabe, wächst. Dieser Trend wird sich noch verstärken, wenn die Techniken besser, preiswerter und auf den vorhandenen Hardwareplattformen lauffähig sind.

Spracherkennung

Von Bedeutung für den Juristen ist hierbei weniger die Eingabe kurzer Befehle zur Programmsteuerung als vielmehr das Diktat längerer Fließtexte, wie er es aus der täglichen Arbeit gewohnt ist. Die Aufgabe, aus einem (mit oder ohne Pausen zwischen den Wörtern) gesprochenen Text eine Folge von korrespondierenden, geschriebenen oder am Bildschirm dargestellten Wörtern zu erzeugen, wollen wir im folgenden als "Spracherkennung" bezeichnen.

Je natürlicher dabei die Eingabe erfolgen kann, also je fließender ein Sprecher sprechen darf, ohne auf Verständnisprobleme beim Computer zu stoßen, desto größer wird die Akzeptanz solcher Systeme sein.

Heutige Probleme

Gerade die Analyse fließend gesprochener Sprache stellt aber ein großes Problem für die heute im Einsatz befindlichen Computer dar. Ein Ausweg wird sich in der fortschreitenden Prozessortechnik bieten, z.B. durch RISC CPUs, aber auch algorithmisch könnten sich andere Ansätze, z.B. neuronale Netze, durchsetzen und das Problem auch auf "kleineren" Maschinen mit entsprechenden Coprozessorkarten zufriedenstellend lösen. Es wäre denkbar, daß in wenigen Jahren die Eingabe längerer Fließtexte über die natürliche Sprache eine selbstverständliche Fähigkeit des Computers sein wird, wie heute z.B. die Tabellenkalkulation. Dies würde den Dialog mit dem Computer, der dann natürlich auch sprachgesteuert funktionieren könnte, erheblich verändern. Es wären andere Oberflächen, z.B. mit natürlichsprachlicher Rückmeldung, und auch andere Anwendungsstrukturen erforderlich.

Textverständnis

Obwohl es so scheint, als solle ein Computer, dem man ein ganzes Urteil diktieren kann, den Befehl: "Setze bitte das Wort Urteil fett." ebenfalls *verstehen* können, statt den Anwender zu Befehlen wie "Cursor hoch Cursor hoch Cursor hoch Markieren ein rechts rechts rechts rechts rechts rechts Zeichen Fett Eingabe" zu zwingen, liegt das Textverstehen als nächster logischer Schritt noch in weiter Ferne. Ohne ein *Textverständnis* ist es dem Computer jedoch nicht möglich, wie der Mensch Annahmen darüber zu machen, welches Wort aus einer Reihe ähnlich klingender gemeint ist.

Dazu ein Beispiel: Der Briefbeginn "Sehr geehrte Damen und Herren" könnte z. B. durch Nebengeräusche, Band-dropouts oder undeutliche Aussprache zu "Sehr geehrte amen und Herren" verstümmelt werden. Dennoch scheint es einem Menschen unmöglich, sich dabei zu verhören und jede Sekretärin würde selbstverständlich "Damen" statt "amen" schreiben, weil sie *weiß*, daß nach "Sehr geehrte" und vor "und Herren" fast kein anderes Wort möglich ist. Ein Computer ohne dieses Wissen könnte nur schreiben, was er gehört hat, und der Satz würde fehlerhaft erkannt.

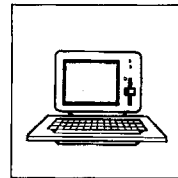
Daß die menschliche Aussprache tatsächlich so undeutlich ist, daß auch Menschen ständig auf solches Erfahrungswissen aus dem Kontext zurückgreifen, wird in Situationen deutlich, in denen kein Kontextbezug möglich ist, also z.B. beim Hören von Namen oder unerwarteten Wörtern z. B. Fremdwörtern oder Fachbegriffen aus fremden Fachbereichen. Auch Menschen sind in solchen Fällen unfähig, die soeben gehörten Wörter niederzuschreiben, d. h. ohne den Kontextbezug sinkt auch beim Menschen die Zuverlässigkeit der Spracherkennung.

Nimmt man hinzu, daß Menschen weitere Informationsquellen über das besprochene Thema, den Kontext, haben, z. B. durch Gestik und Mimik, so wird klar, daß Maschinen, die diese nicht haben, die Spracherkennungsaufgabe nicht hundertprozentig lösen können.

Die Abbildung zwischen dem Klangmuster der Sprache und dem geschriebenen Text ist zudem nicht eindeutig: Ohne Wissen um "Rechtschreibung" ist es nicht entscheidbar, ob das Wort "Gesamtschuldner" mit einem oder zwei "m" geschrieben wird oder ob sich "Sachsen" vielleicht "Saksen" oder gar "Saxen" schreibt. Die Hoffnung auf einen Homunkulus in der Maschine, der "schreibt, wie man's spricht", wird man zurückstellen müssen.

Wörterbücher und Kontextwissen

Jörg H. Blankenburg und Manfred
Weihermüller †, GMD-FIT, Sankt
Augustin.



Es bleibt also, gleich welches Verfahren man zur Spracherkennung wählt, unumgänglich, dem Computer Wortschätze an die Hand zu geben, aus denen er Wörter erkennen soll. Wieviele Fehler pro Seite dem Computer dabei zugebilligt werden, bevor man von einem unbrauchbaren System spricht, ist natürlich ein subjektives Maß, welches, von Tippfehlern abgesehen, wie oben angedeutet, höchstens an "durchschnittlicher" menschlicher Hörleistung gemessen werden darf. In der Justiz, in der es schon heute weniger darauf ankommt, in welcher Qualität ein Urteil den Bürger erreicht, als darauf, daß möglichst viele Urteile von möglichst wenigen Beteiligten produziert werden können, mag diese Schranke niedriger liegen als in der Medizin, wo ein falsch erkanntes Wort des Radiologen das Leben des Patienten kosten könnte. Die Erfahrung zeigt, daß die Trefferquote bei 95 % oder darüber liegen sollte, um das System als brauchbar zu empfinden.

Um eine solche Quote zu erreichen, ist es nötig, dem Computer ein, wenn auch sehr rudimentäres Kontextwissen beizubringen. Dies geschieht in den am Markt befindlichen Systemen durch Training der entsprechenden neuronalen Netze oder Erstellung von Statistiken darüber, mit welcher Wahrscheinlichkeit ein Wort auf ein oder zwei vorhergehende folgt. Es ist umgekehrt ganz erstaunlich, daß ein so rudimentäres Wissen für eine so hohe Trefferquote ausreicht.

Die Größe der Wortschätze ist bei den heutigen Systemen auf 30.000 bis 80.000 Wortformen beschränkt, d. h. jede Deklination, Konjugation, Pluralbildung etc. eines Wortes mit anderer Schreibweise zählt als eigene Wortform. Jeder Sprecher kann dann noch zusätzlich 2.000 bis 5.000 eigene Wortformen eingeben. Es ist angesichts dieser Beschränkung angebracht, nicht nur einen Wortschatz zu benutzen, sondern mehrere fachspezifische, solange man sicher ist, daß der Sprecher während des Diktats wenigstens passagenweise in einem Fachbereich bleiben wird. Im folgenden wird ein Wortschatz zusammen mit dem Kontextwissen als Sprachmodell bezeichnet.

Um beurteilen zu können, ob es überhaupt sinnvoll ist, für einen Fachbereich ein Sprachmodell (bestehend aus Wortformen-schatz und Statistik) zu bilden, ist es also von Interesse zu erfahren, wieviele verschiedene Wortformen benutzt werden (aktiver Wortschatz) und wie häufig diese im Kontext auftauchen. Würde man sehr viele (mehr als 100.000) verschiedene Wortformen finden, die sehr häufig auftreten, so wäre das Fachgebiet zu groß für ein einzelnes, heute nutzbares Sprachmodell. Erfahrungen mit großen publizierten Textmengen aus dem juristischen Bereich lassen solche Befürchtungen aufkommen.

Unser Bestreben war es, ein Sprachmodell für den erstinstanzlichen Richter in allgemeinen Zivilsachen zu erstellen, um die Brauchbarkeit der Spracherkennung an der Stelle zu erproben, wo sie aufgrund des Massengeschäfts am hilfreichsten für die Justiz zu sein scheint. Im folgenden werden also mit "Texten" oder "Daten" in der Hauptsache Gerichtsentscheidungen gemeint sein.

Ein Kriterium, welches sich direkt auf das Sprachmodell auswirkt, ist die Wahl des zugrundeliegenden Textkorpus. Wählt man publizierte Texte, so erfaßt man, z. B. durch einen hohen Anteil höchstrichterlicher und/oder nachbearbeiteter Entscheidungen, eine Sprache, die eventuell von der des erstinstanzlichen Richters erheblich abweicht, so daß dieser mit dem so gebildeten Sprachmodell keine brauchbaren Erkennungsraten mehr erreichen würde. Einerseits erscheinen also die verfügbaren publizierten Daten (Texte) nicht besonders geeignet, andererseits stehen die Texte, die die "Alltagssprache" eines erstinstanzlichen Zivilurteils widerspiegeln, nicht in ausreichender Menge in maschinenlesbarer Form zur Verfügung. Dabei wird die Mindestgröße eines Textkorpus bei ca. 700.000 Wörtern (oder 2000 bis 3000 DIN A4 Seiten oder 4-5 MB ASCII Daten) liegen müssen, um eine halbwegs brauchbare statistische Auswertung zu ermöglichen.

Zur Zeit liegen nur wenige Texte in den Justizbehörden in maschinenlesbarer Form vor. Dies liegt unter anderem daran, daß man in der Justiz die Bedeutung der Information noch nicht erkannt hat und den Wert des Datenmaterials gering einschätzt. So werden PCs in den Schreibkanzleien immer noch nur als bessere Schreibmaschinen angesehen und im hier betrachteten Zusammenhang wertvolle Daten einfach gelöscht. Oft sind es datenschutzrechtliche Bedenken, die dringend geklärt werden müßten, die dazu führen, daß maschinenlesbare Daten "sicherheitshalber" gelöscht werden. Die Archivierung nach vorhergehender Anonymisierung wäre sicher sinnvoller.

Für das hier gewählte System mußten die Daten zur Auswertung außer Haus gegeben werden, und somit, da es sich um sensible personenbezogene Daten handelt, zunächst anonymisiert werden.

Das Maß für die Güte der maschinellen Spracherkennung: die durchschnittliche menschliche Hörleistung

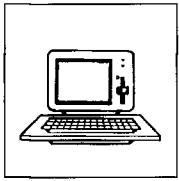
Heutige Systeme

"aktive Wortschätze"

Ein Wortschatz für Zivilrecht

Der zugrundeliegende Textkorpus

Textmangel und andere Probleme



Anonymisierung

Um den Kontext zu erfassen, ist es wichtig, daß die Texte als ganze Sätze vorliegen. In juristischen Texten stehen oftmals Pseudosätze wie z. B.:

IM NAMEN DES VOLKES
URTEIL

In dem Rechtsstreit ...

oder

durch Vernehmung des Zeugen

Herrn Peter Mumpitz, Nixwie Weg 12, 12345 Nixdorf

– vom Kläger benannt –

Es bietet sich an, solche "Sätze", deren grammatikalische Richtigkeit teilweise umstritten ist, die nicht das darstellen, was man unter Fließtext versteht und die bei der Texteingabe schneller über Textbausteine zu handhaben sind als über eine Spracheingabe, insbesondere, weil sie meist noch ungewöhnlich eingerückt sind und viele Anschriften enthalten, bei der Anonymisierung für die Spracherkennung gleich mitzuentfernen. Für die Sprachmodellbildung sind möglichst große Datenmengen aus den wichtigsten Arbeitsbereichen von Interesse, also z. B. auch Beweisbeschlüsse. Bei der Anonymisierung muß hier also eine Vielzahl von Dokumenttypen berücksichtigt werden.

Anonymisierungsqualität

Der Maßstab, der an die Anonymisierungsqualität zu stellen ist, ergibt sich aus der bisherigen Anonymisierungsqualität bei der Veröffentlichung in Büchern oder bereits vorhandenen Datenbanken.

Zu bedenken ist allerdings, daß die Verfügbarkeit der Daten auf Computern einen systematischeren Zugriff gestattet. Dies dürfte sich im Justizbereich aber weit weniger auswirken als auf anderen Gebieten. Aus anonymisierten Gerichtsentscheidungen läßt sich sehr wenig entnehmen, so daß eine Anonymisierung wie bei der Erhebung persönlicher Daten, etwa im Versicherungswesen etc. in der Weise, daß selbst bei rasterfahndungsartigem Zugriff gewisse Rückschlüsse auf Personen nicht mehr möglich sind, weder in der Justiz üblich noch erforderlich ist, u. a. weil die Daten nicht in Datenbankstrukturen vorliegen, sondern Fließtexte sind, die im übrigen bereits öffentlich verkündet wurden.

Unter diesen Gesichtspunkten sollte eine hinreichende Anonymisierung auch schon mit Standardsoftware möglich sein, wenn die Eingangsdaten bestimmte Bedingungen erfüllen.

Es bleibt zu hoffen, daß der Wert der in den Kanzleien erfaßten Daten bald erkannt wird und die Möglichkeiten moderner Standardtextverarbeitungsprogramme zur Standardisierung (z. B. mittels Textbausteinen und Druckformatvorlagen) genutzt werden, damit eine leichtere Anonymisierung und dauerhafte Speicherung zum Zwecke der zulässigen Nutzung möglich wird.

Vorbearbeitung

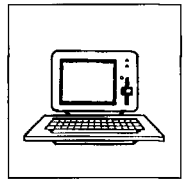
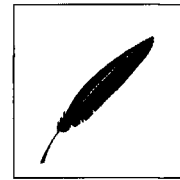
Für unser Sprachmodell war es mit Hilfe einiger Makros für die unterschiedlichen Dokumenttypen leicht möglich, eine grobe Vorbearbeitung der Texte vorzunehmen, z. B. Zeugenanschriften aus Beweisbeschlüssen zu entfernen etc.

Wörterbuchabgleich

Der nächste Schritt bestand darin, ein Wörterbuch (in Form einer Access Datenbank) zu erstellen, in dem keine Namen vorkommen. Dies geschah in Form eines Bootstrapping: Zunächst wurden alle Wortformen des ersten Dokumentes dem Wörterbuch hinzugefügt und anschließend alle Namen manuell entfernt. Dann wurden alle Wortformen des nächsten Dokumentes, die nicht in dem so bereinigten Mini-Wörterbuch standen, in eine separate Liste geschrieben. Aus dieser Liste wurden wieder manuell alle Namen entfernt und die verbleibenden Wortformen dem Wörterbuch hinzugefügt. Die Liste wurde dann gelöscht. So wurde mit allen Dokumenten verfahren. Die entfernten Namen wurden in den Dokumenten automatisch mit Word for Windows durch das Wort [ANONYM] ersetzt. Im Laufe des Verfahrens blieben immer mehr Namen und immer weniger neue Wortformen in der separaten Liste "hängen", so daß die Anonymisierung eines Dokumentes immer schneller ging, wobei das Wörterbuch immer größer wurde. Während für die ersten 4,5 MB noch ca. 5 Tage benötigt wurden, ließ sich das letzte Megabyte einschließlich Vor- und Nachbearbeitung in 4 Stunden anonymisieren.

Nachbearbeitung

Die Nachbearbeitung bestand in der Suche nach Namen in der Umgebung von Schlüsselwörtern wie "Herr", "Frau", "Zeuge", "Firma", "Notar" etc. Dazu wurde das gesamte



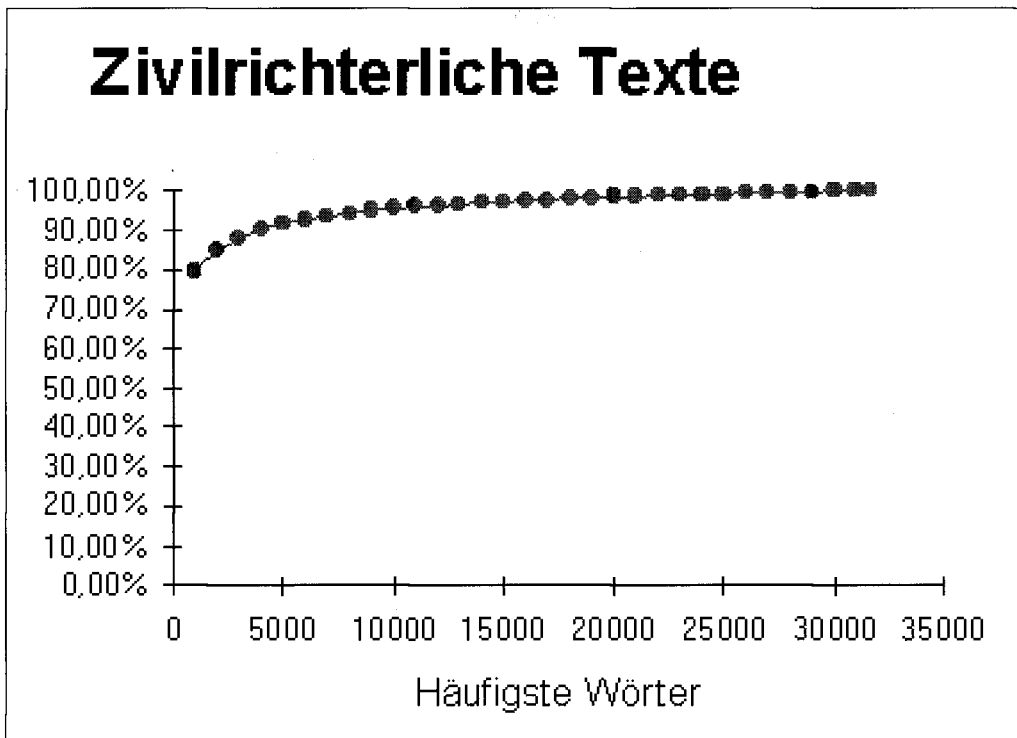
Datenmaterial eines Durchgangs in einer einzigen Datei gehalten, so daß globales Suchen und Ersetzen alle Vorkommen eines Namens in allen Dokumenten erreichen konnte.

Das Ergebnis der Anonymisierung waren in einem Korpus von ca. 800.000 Wörtern 11.000 ersetzte Namen (ohne die in der Vorbearbeitung gelöscht).

In einem letzten Schritt wurden nun die Vorkommen einer jeden Wortform gezählt. Danach bestand der relativ kleine Korpus von 800.000 Wörtern aus ca. 32.000 verschiedenen Wortformen, wobei Groß- und Kleinschreibung unterschieden wurden. Ca. 20.000 Wortformen kamen zweimal oder häufiger vor.

Um die Aussichten für die Erkennungsquote zu beurteilen, ist entscheidend, welcher Anteil des Textkorpus mit wievielen Wortformen abgedeckt wird. Dies zeigt die nachfolgende Abbildung. Bereits mit 1.000 Wortformen sind 80% des gesamten Textkorpus abgedeckt. Mit den 20.000 Wortformen, die zweimal oder häufiger vorkommen, sind 98,5% des gesamten Korpus abgedeckt, eine Zahl, die eine gute Trefferquote für die Spracherkennung erhoffen läßt.

Ergebnisse



Die mit diesem Sprachmodell erzielten ersten Testergebnisse sind schwer einzuordnen. Beim Diktieren können zwei Arten von Fehlern auftreten: Zum einen die "Hörfehler", die vorliegen, wenn der Rechner ein ihm *bekanntes* Wort nicht "versteh", d. h. ein anderes Wort schreibt als gesprochen wurde, obwohl das gesprochene im Sprachmodell enthalten ist. Dieser Fehler wiegt schwer und seine Anzahl sollte man ins Verhältnis zur diktierten Gesamtwortzahl setzen, um die Erkennungsrate zu ermitteln. Der zweite Fehler ist der "Wortschatzfehler", der vorliegt, wenn ein anderes Wort geschrieben wurde, weil das gesprochene Wort nicht im Sprachmodell vorhanden war. Im Fall des Wortschatzfehlers kann der Computer das Wort nachträglich erlernen.

Durch alle Tests hindurch traten etwa gleich viele Fehler beider Arten auf. Insbesondere die Wortschatzfehler sollten sich bei zunehmender Testdauer verringern, weil die sprecher-spezifischen Wörter irgendwann zum größten Teil erlernt sein sollten.

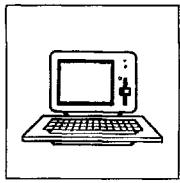
Die Fehlerrate hängt allerdings von vielen weiteren Faktoren ab, die hier angerissen werden sollen und die eine Einordnung erschweren und weitere Tests erforderlich machen.

Die erste Testaufgabe bestand darin, dem System Texte vorzulesen, die in die Erstellung des Sprachmodells eingegangen waren. Soweit hierbei Wortschatzfehler auftraten, handelte es sich um die ca. 12.000 Wörter, die nur einmal im gesamten Korpus vorkamen. Diese sind offensichtlich gleichmäßig über die gesamten Texte verteilt, denn ca. 3 % aller Fehler aus den Textstichproben waren solche Wortschatzfehler. Mit etwa ebenso vielen Hörfehlern ergab sich eine Erkennungsquote von 97 %, wenn man nur die Hörfehler zählt, und von immer noch 94 % wenn man beide Fehlerarten zählt.

Erste Tests

Mögliche Fehlerursachen

Bekannte Texte



Etwa das gleiche Bild ergibt sich, wenn man zivilrichterliche Texte diktiert, die nicht in die Sprachmodellbildung eingingen. Diese Tatsache läßt darauf hoffen, daß das Sprachmodell für den Echteininsatz sehr brauchbar ist.

Leider gab es auch immer wieder Ausreißer, d. h. solche Texte, mit denen die Erkennungsrate deutlich unter 90 % lag. Es wurden auch nicht genügend Texte vorgelesen, um hier zu wirklich statistisch relevanten Aussagen zu kommen. Der Grund dafür ist der, daß der subjektive Brauchbarkeitsmaßstab im Echteininsatz von anderen, im folgenden näher besprochenen Faktoren stärker beeinflußt zu sein scheint.

Einschränkungen

Diktierweise

Weniger Probleme als erwartet macht es, sich an die abgehackte Sprechweise bei dem (zum Zeitpunkt der Drucklegung) unter dem Namen "Voice Type" für Windows und OS/2 verfügbaren IBM-System zu gewöhnen. Allerdings ist dies auch wieder stark vom Benutzer abhängig.

Eigennamen

Erhebliche Probleme bereiten Eigennamen, die das System nicht kennen kann. Es gibt mehrere Strategien, damit umzugehen. Zum einen kann man die Namen dem System sofort beim ersten Auftreten oder gar vor dem Diktieren beibringen und dann später flüssig weiterdiktieren. Zum anderen kann man eine Zuordnungsliste von dem System bekannten Wörtern zu Eigennamen (z. B. Eichhörnchen, Marder, Fuchs bedeuten Herr Maier, Herr Müller, Firma Schmitz GmbH) haben. Die dem System bekannten Wörter sollten sich klanglich unterscheiden, um Verwechslungen, die der Korrekturleser nur schwer bemerken würde, zu vermeiden, also (Zeugeeins, Zeugezwei, Zeugedrei wären daher ungeeignet). Beim Diktieren stockt man beim Blick in die Zuordnungsliste und vergißt womöglich den Rest des Satzes. Ignoriert man das Namensproblem und diktiert einfach die unbekannt Namen, sinkt die Erkennungsquote, weil der Rechner anstelle der Namen irgendwelche ähnlichklingenden Wörter verwendet, auf die aber andere Wörter erwartet werden, als wirklich gesprochen wurden, weshalb dann Folgefehler auftreten können.

Korrekturen

Das Korrekturlesen der diktierten Texte ist erheblich schwieriger als das von gewöhnlichen Bandabschriften, denn diese Diktate enthalten ausschließlich korrekt geschriebene Wörter, deren Fehlerhaftigkeit man nur am mangelnden Sinn erkennt, den man wiederum nur kennt, wenn man um das Diktierte weiß. Zum Beispiel ist für einen Außenstehenden nicht erkennbar, ob der Satz "Er erhielt den Auszahlungsbetrag." nicht heißen müßte: "Er erhielt den Anzahlungsbetrag.", obwohl der Unterschied erheblich sein kann. Eine (zeitaufwendige) Möglichkeit ist es deshalb, das Diktat, welches im Rechner mitgeschnitten wird, erneut abzuhören, gleichzeitig mitzulesen und Fehler zu korrigieren.

Fremde Sachgebiete

Es kommt ebenfalls immer wieder vor, daß ein Sachverhalt, ähnlich wie Eigennamen, viele Begriffe aus einem Sachgebiet beinhaltet, welche nicht im Sprachmodell enthalten sind, z. B. eine Klage, in der viele rentenversicherungstechnische Ausdrücke vorkommen. Auch hier empfiehlt es sich, die Ausdrücke dem System sofort beizubringen, um die dann ohnehin nicht mehr sehr hohe Erkennungsrate nicht noch zu verschlechtern.

Aus all dem ergibt sich, daß eine absatzorientierte Arbeitsweise am effizientesten ist. Der Sprecher sollte seinen Absatz diktieren, ohne auf dem Bildschirm zu verfolgen, was der Rechner daraus macht, schon weil die dort erscheinenden Wörter eventuell wieder revidiert werden, es also keinen Sinn macht, zu verfolgen, ob der Rechner den Sprecher "richtig verstanden" hat. Dann werden der Absatz markiert und das Gesprochene erneut abgehört oder bei kürzeren Absätzen die Fehler korrigiert. Dazu ist es sehr oft notwendig, ganze Wörter neu einzutippen. Auch bei Hörfehlern stehen die richtigen Wörter, die dem Rechner ja bekannt sind, oft nicht in der Liste der Ähnlichklingenden, so daß man auch sie neu eingeben muß. Ein ganz entscheidender Nachteil ist, daß Wortkomposita, wie z. B. "Rentenversicherungsnehmer", oft bereits richtig erkannt werden und als "Renten Versicherung Nehmer" im Text stehen, so daß man eigentlich eine Möglichkeit haben müßte, alle drei Wörter zu markieren und mit einem Mausclick zum Kompositum zu erklären, welches der Rechner dann automatisch lernt (die Aussprache liegt ihm ja ebenfalls vor). Statt dessen ist man gezwungen, das ganze Wort in einer komplizierten Prozedur erneut einzugeben. Noch schöner wäre natürlich, wenn der Rechner die ihm ja ohnehin bekannten und auch richtig erkannten Wörter anhand eines Kompositalexikons gleich richtig setzen würde. Dies würde gerade im juristischen Bereich die Erkennungsrate und das subjektive Brauchbarkeitsgefühl merklich verbessern.

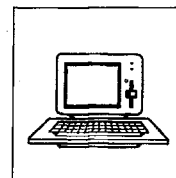
Das getestete Computer-Diktiersystem birgt heutzutage noch einen kleinen Widerspruch in sich: Einerseits ist es von seiner Anwendung her so stark computerorientiert, daß der eingefleischte computerungeübte Diktiergerätebenutzer damit schwer zurechtkommt und man bei ihm eher auf Ablehnung stößt. Andererseits hat der spielend leicht damit umgehende, computergewöhnte Benutzer oftmals das Gefühl, er hätte seinen Text schneller mit der Hand getippt.

Für beide Benutzergruppen ist das System zur Zeit ungeeignet. Das System bietet demjenigen eine erhebliche Beschleunigung seiner Arbeit, der im Umgang mit modernen grafischen Benutzeroberflächen und Editoren firm, beim Tippen von Langtexten aber auf eine nicht immer oder nicht schnell genug verfügbare Schreibkraft angewiesen ist.

Leider ist dieser Benutzertyp noch recht selten, es finden sich eher die zuerst genannten Extreme. Dies dürfte sich aber mit dem weiteren Fortschreiten des Computereinsatzes in der Justiz ändern.

Spracherkennungssysteme machen Fortschritte in Richtung auf kontinuierliche Sprechweise (heute schon beim Philips-System Wirklichkeit) und lassen sich in Ihrer Bedienung dem Diktiergerät stärker anpassen. Der Nachteil, daß die Übersetzung in einem zeitaufwendigen Schritt nach dem Gesamtdiktat erfolgt, was die Fehler durch die beschriebenen Effekte, z. B. bei Namen, Fachbegriffen etc., im Vergleich zur oben beschriebenen absatzweisen Diktiermethode erhöht, wird sich durch schnellere Hardware und größere Wortschätze mildern.

Eine gute Schreibkraft werden solche Systeme nie ersetzen, aber sie werden gut genug werden, um diesen Beruf aussterben zu lassen.



Benutzergewohnheiten

Zukunftsaussichten