

Eingabe strukturierter Befunddaten mittels Spracherkennung: Ein Feldtest

Alexandra Klein¹, Christina Drexel², Dian Tan², Johannes Matiassek¹,

Ralf Birkemeyer³, Harald Trost⁴

Österreichisches Forschungsinstitut für Artificial Intelligence (OFAI), Wien¹

Nuance Communications Austria GmbH, Wien²

Innere Medizin III, Schwarzwald-Baar-Klinikum Villingen-Schwenningen GmbH³

Institut für medizinische Kybernetik und Artificial Intelligence, Zentrum für Hirnforschung,
Medizinische Universität Wien⁴

Zusammenfassung

Dieses Papier beschreibt einen Feldtest, der durchgeführt wurde, um in einer eingeschränkten Domäne Erkenntnisse zur Benutzbarkeit eines Spracherkennungssystems für die Eingabe von strukturierten Daten aus dem medizinischen Bereich zu gewinnen. Bei der Eingabe von medizinischen Befunden wird in Krankenhäusern zunehmend vom klassischen Diktieren, bei dem ein ganzer Bericht in einem Zug diktiert wird, zur interaktiven Eingabe von strukturierten Daten übergegangen, was eine Umstellung für das medizinische Personal bedeutet. Die strukturierte Eingabe bietet den Vorteil, dass die Daten leicht für andere Zwecke, wie beispielsweise Dokumentation, Abrechnung und Qualitätskontrolle, verwendet werden können. Da sich der Einsatz von Spracherkennungstechnologie in Diktierszenarien mit eingeschränktem Wortschatz, z.B. in der Radiologie, bewährt hat, ist zu erwarten, dass Spracherkennung auch die Eingabe von strukturierten Daten erleichtern kann. Um den Ansatz in einem echten Arbeitsumfeld zu erproben, wurde ein Prototyp für die EKG-Befundung entwickelt und mit Kardiologen in einem Krankenhaus getestet. Die Ergebnisse der Feldtests sind ermutigend: Auswertungen von Logfile-Daten, Fragebögen, Interviews mit den Benutzern und eine Analyse des Äußerungs-Korpus haben ergeben, dass die Benutzer mit der Spracherkennung und den Regeln der Eingabesprache strukturierte Daten eingeben können und bei der Erledigung ihrer alltäglichen Aufgaben davon profitieren.

1 Einleitung

Krankenhäuser benötigen Patientendaten zunehmend in strukturierter Form, um sie für Dokumentation, Abrechnung und Qualitätskontrolle zu verwenden. Medizinisches Personal muss die Informationen daher interaktiv in Datenbanken eingeben. Dies bedeutet eine Umstellung im Arbeitsablauf, da Berichte vorher meist von Ärzten als freier Text diktiert, dann von Schreibkräften transkribiert und schließlich von Ärzten kontrolliert wurden. Das neue interaktive Paradigma eröffnet Potenzial und Notwendigkeit für Methoden zur effizienten und einfachen Dateneingabe. Es ist zu erwarten, dass Spracherkennung die Eingabe erleichtern kann, z.B. wenn die Verwendung von Tastatur und Maus unpraktisch ist, da Ärzte und Pfleger gleichzeitig Ausdrucke von Befunden oder Notizen in der Hand halten.

Es gibt zwei mögliche Vorgehensweisen, Spracherkennung in die strukturierte Dateneingabe einzubinden. Eine Möglichkeit besteht darin, ein vorhandenes GUI, das bisher mit getippten Eingaben und Mausklicks bedient wird, durch Sprachbefehle ansteuerbar zu machen. Hier erhofft man sich einen Effizienzgewinn für die Benutzer durch die Spracheingabe. Die zweite – weit aufwändigere – Möglichkeit ist die Entwicklung eines Systems zur strukturierten Dateneingabe in einem Umfeld, in dem bisher mit Freitext gearbeitet wurde. Hier ist die Umstellung für die Benutzer größer, da sich nicht nur die Eingabemodalität ändert, sondern generell die Arbeitsweise, indem statt freiem Text strukturierte Daten eingegeben werden müssen. Auch die Entwicklung eines solchen Systems ist anspruchsvoller, da 1) die Struktur der einzugebenen Daten erst modelliert werden muss, 2) mögliche Interaktionen mit dem System spezifiziert werden müssen und 3) ein Interface-Design erstellt werden muss. Durch den zweiten Ansatz ist jedoch ein größerer Effizienzgewinn, nicht nur in Bezug auf die Spracheingabe, sondern auch auf die Strukturierung und Wiederverwertbarkeit der Daten, zu erwarten, die vorher nur in Form von Freitext vorlagen. Im folgenden soll die Realisierung und Erprobung eines solchen Systems für ein eingeschränktes Anwendungsgebiet, in unserem Fall EKG-Befundung, beschrieben werden.

2 Strukturierte Eingabe mit Spracherkennung im medizinischen Bereich

Bei der Eingabe von strukturierten Daten mittels Spracherkennung handelt es sich um eine neuartige Technologie. Entsprechend mangelt es bisher an Studien zu Benutzerfreundlichkeit und Akzeptanz der Systeme. Zwar wurde schon in den 90er Jahren auf das Potenzial der Kombination von Spracherkennung und strukturierter Eingabe verwiesen (Wormek et al. 1997; Teel et al. 1998), da einerseits Spracherkennungstechnologie in Freitextanwendungen – besonders in der Radiologie - erfolgreich verwendet wurde, andererseits aber der Bedarf nach strukturierten, wiederverwertbaren Daten zu wachsen begann. Doch konnten sich solche Systeme zunächst nicht durchsetzen, was hauptsächlich auf die damalige Qualität der Spracherkennung zurückzuführen ist (Grasso 2003). Spracherkennung wird daher weiterhin

fast ausschließlich in Diktieranwendungen verwendet, bei denen die Texte von medizinischen Transkriptionisten nachbearbeitet werden und die Benutzer somit nicht direkt mit dem Spracherkennungsergebnis konfrontiert sind. Interaktive Systeme verwenden häufig Textbausteine, die auf eine eingeschränkte Anwendung zugeschnitten sind und die direkt vom Benutzer mit Maus oder Tastatur ausgewählt werden müssen. Falls keiner der Textbausteine der Benutzerintention entspricht, muss entweder ein anderer Textbaustein ausgewählt werden oder eine Freitexteingabe per Tastatur gemacht werden (Bell et al. 1992).

Auch in anderen medizinischen Bereichen gibt es wenig Erkenntnisse zum Einsatz von interaktiven Systemen mit Spracherkennung. (Grasso 2002) stellt ein System zur gesprochenen Eingabe für klinische Daten aus Studien vor. (Starlander et al. 2005) beschreiben die Evaluierung eines Systems zum maschinellen Dolmetschen von Fragen, die Ärzte (bzw. in der Evaluierung Medizinstudenten) an Patienten richten. Nach drei Testsitzungen gelang es einigen Teilnehmern, das System ebenso gut zu bedienen wie die Entwickler des Systems.

Seit einiger Zeit gibt es wieder Versuche, Spracherkennung als Eingabemodalität für strukturierten Daten zu erproben. Begünstigende Faktoren sind die verbesserte Spracherkennungsqualität und die Tendenz zu standardisierten elektronischen Patientenakten (Tang & McDonald 2001; Kauppinen et al. 2008). Während erste Systeme zur strukturierten Eingabe mit Spracherkennung in eingeschränkten Anwendungen beschrieben werden (z.B. Nagy et al. 2008), gibt es bisher kaum Erkenntnisse zur Benutzbarkeit dieser einzelnen Systeme im Arbeitsalltag. Zur Untersuchung der allgemeinen Akzeptanz von Spracherkennung im Krankenhausumfeld haben (Alapetite et al. 2009) Benutzern von Spracherkennungstechnologie in einem dänischen Krankenhaus – das erste in Dänemark, das Spracherkennung auf allen Stationen eingeführt hat – zunächst zu Beginn und später nach einiger Zeit der Benutzung Fragebögen vorgelegt. Dabei wurde festgestellt, dass die Realität häufig nicht den Erwartungen entsprach. Die befragten Benutzer schätzen allerdings besonders die Tatsache, dass die Berichte unmittelbar zur Verfügung stehen, was bei dem Diktierszenario mit anschließender Transkription naturgemäß nicht der Fall ist.

3 EKG-Prototyp

Das EKG ist eine nicht-invasive Prozedur, die mindestens einmal für jeden Patienten auf einer Kardiologie-Station durchgeführt wird. Bei der Befundung wird das Ergebnis dokumentiert und interpretiert, bisher in der Regel in Freitext. Spracherkennung kommt für diesen Anwendungsfall mit einem geringen Vokabular und eingeschränkten Formulierungen aus. Ein Domänenmodell und ein Prototyp des Systems konnten daher in kurzer Zeit fertiggestellt werden. Allerdings ist in dem eingeschränkten Anwendungsfall der EKG-Befundung durch die strukturierte Eingabe mit gesprochener Sprache weniger Effizienzgewinn zu erwarten als in anderen, komplexeren Anwendungsfällen.

Mit dem Prototyp konnten die Ärzte den EKG-Befund direkt eingeben und später ohne Änderungen für einen formalen Bericht verwenden. Um die Vorteile des Systems voll zu nutzen, mussten sich die Benutzer mit der Bedienung des Systems und den Regeln der Eingabesprache vertraut machen. Die Eingabesprache wurde dabei möglichst einfach und das GUI möglichst übersichtlich gestaltet. Gesprochene Äußerungen können nicht nur für Eingaben, sondern auch für Navigations- und Korrekturbefehle verwendet werden.

Da in unserem Fall - im Gegensatz zu der von (Alapetite et al. 2009) beschriebenen Untersuchung – Endbenutzer am Entwicklungsprozess beteiligt waren, konnten ihre Anforderungen und Präferenzen bei der Entwicklung des Systems berücksichtigt werden. Das gilt insbesondere für die Entwicklung des Domänenmodells und des GUI, welches das Domänenmodell darstellt. Die erste Version des Domänenmodells wurde von einem Kardiologen entworfen. Eine optimierte, von zwei Kardiologen getestete Version mit 991 Knoten wurde im Feldtest verwendet. Das Domänenmodell ist an den Benutzererfordernissen orientiert, jedoch lässt sich das System leicht an Veränderungen am Domänenmodell, falls von den Benutzern gewünscht, und neue Domänenmodelle und Anwendungsszenarien anpassen.

Bei unserem System wird Semantik der Spracherkennung zugrunde gelegt, indem alle Benutzereingaben auf die Domänenmodell-Repräsentation abgebildet werden. Diese semantische Repräsentation ist mit einer eingeschränkten (kontrollierten) Eingabesprache verbunden, um aus der Sicht von Semantik und Spracherkennung die Ambiguität und Komplexität der Äußerungen zu reduzieren. Alle Eingaben sind interaktiv. Unerfahrene Ärzte profitieren besonders von der durch die Anzeige des Domänenmodells im GUI vorgegebenen Berichtsstruktur. Fehler bei der Erkennung und Interpretation von Äußerungen müssen von den Benutzern selbst korrigiert werden.

3.1 GUI

Abbildung 1 zeigt den Bildschirm, nachdem einige Eingaben verarbeitet worden sind. Das Menü erlaubt die Verwaltung von Dateien, Ansichten und Benutzerkonten, ein Benutzerleitfaden steht zur Verfügung. Die wichtigsten Operationen (neu/öffnen/speichern) sind mit Buttons im Tool-Balken zu steuern. Auf der linken Seite des GUI ist die Hierarchie im Domänenmodell dargestellt. Bewegt man die Maus über ein Konzept, so erscheinen Hinweise zur Eingabesprache für den lokalen Kontext als Tooltips. Auf der unteren rechten Seite ist die Formular-Ansicht des Unterbaums im derzeitigen Fokus mit Checkboxes, Radio Buttons und Eingabefeldern zu sehen. Oben rechts in der Zusammenfassung können die Benutzer ihre Eingaben mitverfolgen und kontrollieren. Weitere Details werden als Links dargestellt (z.B. *P-Welle*). Rechts oben über der Formular-Ansicht ist die Zusammenfassung der eingegebenen Daten zu sehen. Es ist möglich, durch Klicken auf die Datenelemente zu diesem Punkt zu navigieren; dabei werden die Baum- und Formularansichten entsprechend synchronisiert. Der Text kann mittels des Kontext-Menüs in verschiedenen Formaten exportiert werden.

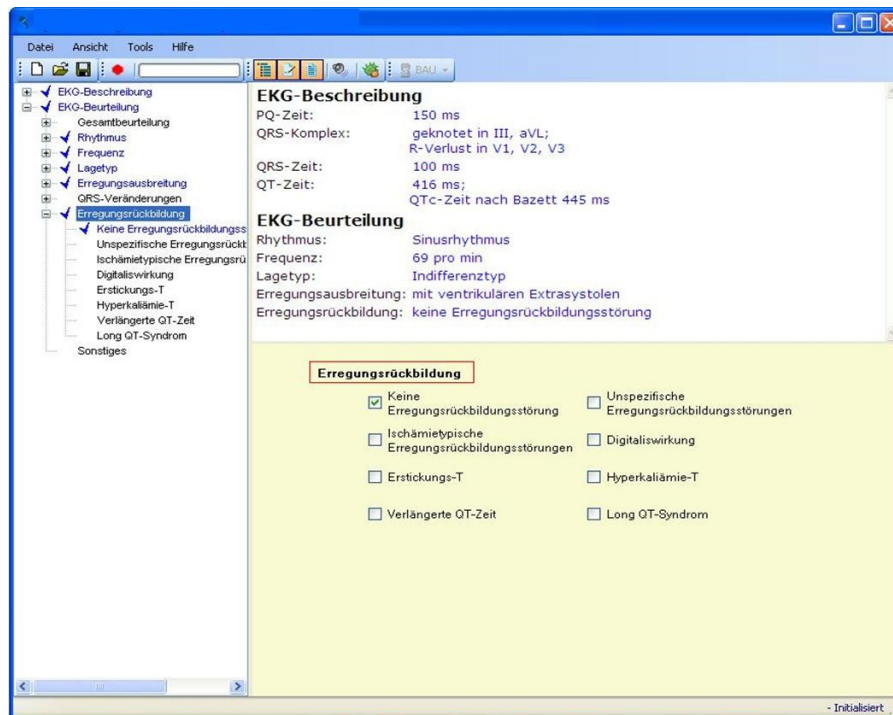


Abbildung 1: Benutzeroberfläche mit eingegebenen Daten

3.2 Eingabesprache

In der Regel sprechen Benutzer die Überschriften der Abschnitte (z.B. *Beurteilung* oder *Beschreibung*) oder Unterabschnitte (z.B. *P-Welle*) und Werte. Es ist möglich, Kurvenbeschreibungen mit Attributen und Ableitungen sowie Zahleneingaben (also Messwerte) und Befundungen im Beurteilungsteil einzugeben. Bei Kurvenbeschreibungen können die Attribute vor oder nach dem Kurvenelement stehen, z.B. *T-Welle flach in römisch eins zwei und aVL* oder *spitze T-Welle in den Ableitungen V5 und V6*. Zahlenwerte werden mit einem Bezeichner, der Zahl und optional der Einheit eingegeben, z.B. *PQ-Zeit hundert zwanzig* für 'PQ-Zeit 120 ms'.

Erlaubte Reparaturkommandos sind *Rückgängig* und *Lösche*. *Rückgängig* bezieht sich jeweils auf die zuletzt durchgeführte Aktion. *Lösche* in Kombination mit einem Abschnitt oder Unterabschnitt im Befund löscht diesen. Es ist auch möglich, Werte einfach neu einzugeben, um inkorrekte Eingaben zu überschreiben. Dem Benutzer wird im GUI angezeigt, welche Eingaben vom System erkannt und welchen Elementen im Domänenmodell zugeordnet worden sind. Bei Fehlern muss eine entsprechende Korrektur durch den Benutzer vorgenommen werden.

4 Feldtest

Der Prototyp wurde für Kardiologen entwickelt, und eine Evaluierung im realen Arbeitsalltag der Experten wurde als unerlässlich angesehen, um aussagekräftige Ergebnisse zu erhalten. Zu diesem Zweck wurde ein Feldtest durchgeführt. Von 12 Kardiologen einer Station nahmen 10 an einem Workshop teil, bei dem das Spracherkennungssystem vorgestellt wurde, die Regeln der Eingabesprache erklärt wurden und die Funktionalität des Systems demonstriert wurde. Die Teilnehmer beantworteten einen kurzen Fragebogen zu ihren Computer-Erfahrungen und zu ihrer Expertise in der EKG-Domäne. Eine Testphase von 10 Wochen folgte, während der die Testpersonen so viele EKG-Berichte wie möglich mit dem Prototyp erstellen sollten. Letztlich testeten 7 Kardiologen das System und produzierten damit 172 Berichte.

Es wurden Logfiles aller Aktionen und Aufzeichnungen aller Äußerungen erstellt. Zusätzlich bewerteten die Kardiologen nach jedem Befund die Erkennungsqualität und beantworteten nach der Testphase einen Fragebogen bezüglich ihrer Erfahrungen mit dem System. Mit fünf Teilnehmern wurden zusätzlich Interviews geführt. Wegen der kleinen Stichprobe konnte keine umfassende statistische Analyse durchgeführt werden, doch werden einige Korrelationen aufgezeigt, die als Grundlage für weitere Forschungsarbeiten dienen.

4.1 Charakteristika der Interaktionen

72 aufgezeichnete Berichte wurden manuell transkribiert und mit Informationen zur (angenommenen) Benutzerintention, der Konsequenz der Aktion und (im Fall von Fehlern) der Fehlerquelle annotiert, um zu bestimmen, bei welchen Interaktionstypen am häufigsten Fehler auftraten (Oulasvirta et al. 2006). In den 72 Berichten wurden insgesamt 2.164 Aktionen annotiert. 91,2% (oder 1.974 Aktionen) waren gesprochene Interaktionen, 8,8% waren Interaktionen mit Hilfe von Maus oder Tastatur. Tabelle 1 zeigt die Interaktionstypen und die Benutzerintentionen. 'keine Intention' bezieht sich auf Fälle, bei denen das Spracherkennungssystem beispielsweise aus Hintergrundgeräuschen eine Äußerung erkannte.

	Gesprochen		Maus/ Tastatur	
	#	%	#	%
Intentionen				
Dateneingabe	846	42,9	44	23,2
Navigation	374	18,9	72	37,9
Reparatur	441	22,3	74	38,9
keine Intention	290	14,7		
Intention unklar	23	1,2		

Tabelle 1: Aktionen nach Benutzerintention

Tabelle 2 zeigt einen Ausschnitt aus einer Befundungsinteraktion. Der Benutzer möchte die Eingabe zu Erregungsrückbildung löschen, verwendet aber zunächst nicht wie vorgesehen

lösche Erregungsrückbildung, so dass der Löschbefehl vom System nicht einem eingegeben Wert zugeordnet werden kann. Der Benutzer versucht es schließlich mit der höflichen Form *bitte löschen*, was zu Erkennungsproblemen führt. Nach dem Versuch, explizit zu der zu löschenden Eingabe hinzunavigieren, ist die Reparatur erst bei Gebrauch des richtigen Befehls mit korrekter Syntax erfolgreich, das System kann die Reparatur einem Eintrag zuordnen. Der Benutzer fährt mit weiteren Dateneingaben fort.

gesprochen (Transkription)	Erkennungsergebnis	Benutzerintention
Erregungsrückbildung löschen bitte löschen Beurteilung	Erregungsrückbildung lösche wiederherstellen Beurteilung	1. lösche 'Erregungsrückbildung'
Erregungsrückbildung lösche Erregungsrückbildung Beurteilung	Erregungsrückbildung lösche Erregungsrückbildung Beurteilung	2. 'Beurteilung' eingeben
Frequenz siebzig pro Minute	Frequenz siebzig pro Minute	3. (Herz-)Frequenz' eingeben 4. Wert für 'Frequenz' eingeben
Erregungsrückbildung Hyperkaliämie-T	Erregungsrückbildung Hyperkaliämie-T	5. 'Erregungsrückbildung' eingeben 6. Befund eingeben

Tabelle 2: Beispiel für Interaktionen mit Korrektur und Dateneingabe

Gesprochene Kommandos zur Dateneingabe waren in 62% der Fälle erfolgreich. 10% der Aktionen führten zu einem Rückschritt im Systemzustand. Die meisten nicht erfolgreichen gesprochenen Eingaben gingen wenigstens zum Teil auf Benutzerverhalten zurück, was darauf hindeutet, dass die Benutzer Probleme mit der Eingabesprache hatten. Weniger als 50% der gesprochenen Reparatur-Kommandos waren erfolgreich, d.h. die Reparatur-Intention der Benutzer war bei der Evaluation erkennbar, jedoch benutzten sie nicht auf Anhieb das korrekte Korrekturkommando und benötigten daher mehrere Interaktionen. Dies führt u.a. zu einer hohen Anzahl von Interaktionen, die mit keinem Interaktionsfortschritt verbunden war. Hier ist eindeutig Verbesserungsbedarf im Hinblick auf eine intuitivere Reparaturfunktionalität gegeben. Sprachliche Äußerungen dienen meist der Dateneingabe. Maus-/Tastatur-Aktionen werden für Navigation benutzt. Auf Maus-/Tastatur-Reparaturen wurde zurückgegriffen, wenn Sprachbefehle zur Korrektur nicht erfolgreich waren oder mehrere Interaktionsschritte erfordert hätten.

Als groben Indikator für Effizienzgewinn ziehen wir die Entwicklung des Zeitbedarfs der Äußerungen pro Bericht heran (Abbildung 2):

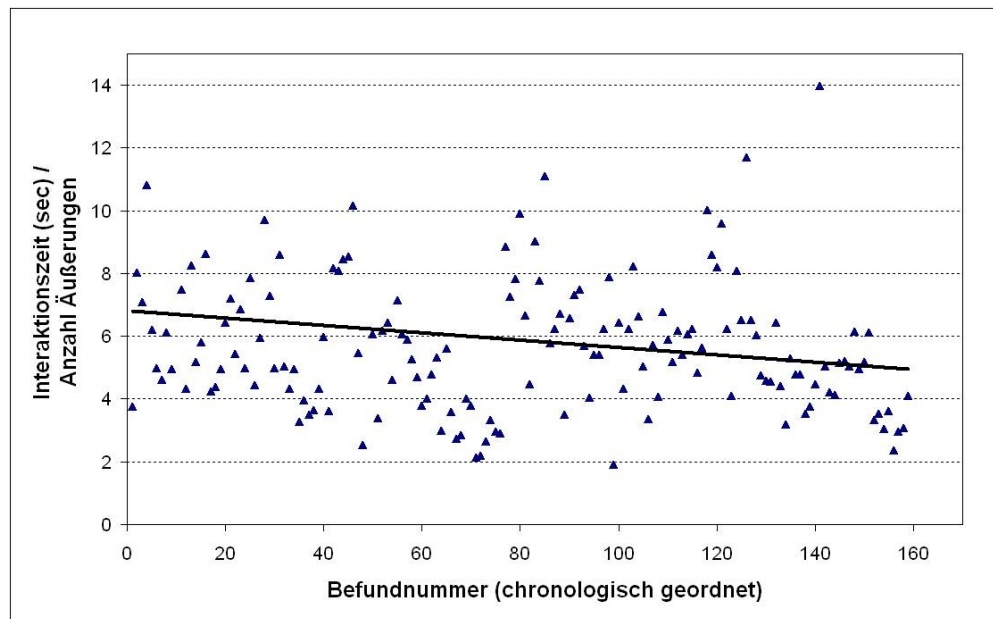


Abbildung 2: Äußerungen bezogen auf die Zeit

In Abbildung 2 repräsentieren Punkte den durchschnittlichen (zeitlichen) Bedarf von Äußerungen pro Bericht (d.h. die Länge des Audio-Signals dividiert durch die Anzahl der Äußerungen). Die Berichte sind chronologisch auf der x-Achse aufgelistet. Es ist ersichtlich, dass sich der Wert allmählich verringert. Dies deutet darauf hin, dass die Benutzer weniger Zeit benötigen, um Äußerungen zu formulieren und so das System effizienter nutzen. Weniger erfahrene Kardiologen profitierten von der strukturierten Eingabe mit Spracherkennung zeitlich mehr; die Anzahl der analysierten EKG pro Woche ist negativ korreliert mit der Zeitersparnis ($r = -.77$). Insgesamt dauerte es jedoch länger, einen Bericht mit Hilfe der Spracherkennung zu erstellen als das übliche Vorgehen, bei dem handschriftliche Notizen erstellt wurden. Wir gehen davon aus, dass dieser Effekt zumindest zum Teil mit der relativ kurzen Testphase und dem wenig umfangreichen Anwendungsfall erklärt werden kann.

4.2 Benutzereindruck

Generell gaben die Benutzer an, von der Spracherkennung einen sehr guten Eindruck zu haben. Die subjektiven Daten beschreiben das System als leicht zu verstehen und zu benutzen. Kleinere Probleme wurden schnell überwunden, und die Benutzer rechnen mit einer weiteren Gewöhnung an das System im Lauf der Zeit. Die Domänenmodell-Baumstruktur wurde von allen Benutzern häufig zur Orientierung konsultiert. Die Zusammenfassung wurde während des ganzen Eingabevorgangs zur Überprüfung der

Erkennungsergebnisse genutzt. Eine effiziente Verwendung der Eingabesprache scheint mit Erfahrung bezüglich des Systems und der Domäne zu korrelieren. Kardiologen mit längerer Erfahrung im Erstellen von EKG-Befunden bewerten subjektiv den richtigen Einsatz der Eingabesprache als weniger problematisch ($r=.51$). Die Antwortzeiten des Systems wurden als gut bewertet.

5 Schlussbemerkung

Die Feldtests lieferten Erkenntnisse zur Benutzerfreundlichkeit von Spracherkennung beim Erstellen von strukturierten Befundtexten. Während des Entwicklungsprozesses und des Feldtests wurde deutlich, wie wichtig es für unseren Ansatz ist, die Benutzer in die Erstellung des Domänenmodells, die Spezifikation des Vokabulars der Eingabesprache sowie der Darstellung des gewünschten Ergebnis-Berichts einzubinden. Für die Entwicklung weiterer Systeme muss allerdings ein Konzept entwickelt werden, wie eine solche Benutzereinbindung systematisch erfolgen kann, indem z.B. Benutzer Tools und kooperative Prozesse bei der Spezifikation des Domänenmodells verwenden. Zudem sind grundlegende Verbesserungen in der Darstellung des Domänenmodells und des eingegebenen Berichts im GUI nötig.

Was die Eingabesprache betrifft, so haben unsere Tests gezeigt, dass die Benutzer in der Lage waren, die eingeschränkte Eingabesprache zu erlernen und zu verwenden. In geringerem Umfang gilt dies für Reparaturmechanismen; hier gibt es Verbesserungsbedarf bei Benutzerfreundlichkeit und Benutzerführung. Insgesamt produzierten die Kardiologen Äußerungen, die effizient in Berichte umgewandelt werden konnten. Wir gehen davon aus, dass der Einsatz von Spracherkennung die Eingabe von strukturierten Daten beschleunigen wird, wobei zu erwarten ist, dass die Vorteile des Ansatzes im Hinblick auf einen Zeit- und Effizienzgewinn in komplexeren Domänen deutlicher hervortreten. Als Testanwendung war die EKG-Befundung jedenfalls sehr hilfreich. Basierend auf den Erfahrungen des Feldtests in der EKG-Domäne wird der Ansatz derzeit auf komplexere Domänen übertragen. Zudem wird an der Entwicklung von Mechanismen zur Bewertung von auf Domänen- und Weltwissen beruhenden Plausibilitäten für Erkennungs- und Interpretationsergebnisse gearbeitet.

Literaturverzeichnis

- Alapetite, A., Andersen, H.B. & Hertzumb, M. (2009). Acceptance of speech recognition by physicians: A survey of expectations, experiences, and social influence. *IJHCS* 67(1), 36-49.
- Bell, D.S., Greenes, R.A. & Doubilet, P. (1992). Form-Based Clinical Input from a Structured Vocabulary: Initial Application in Ultrasound reporting. In: *Proceedings of the Annual Symposium on Computer Applications in Medical Care*, 189-790.

- Grasso, M.A. (2002). Structured speech input for clinical data collection. In: *Proceedings of the 15th IEEE Symposium on Computer-Based Medical Systems*. 199-204 .
- Grasso, M.A. (2003). The Long-Term Adoption of Speech Recognition in Medical Applications. In: *Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems (CBMS)*, 257-262.
- Kauppinen, T., Koivikko, M.P. & Ahovuuo, J. (2008). Improvement of Report Workflow and Productivity Using Speech Recognition: A Follow-up Study. *Journal of Dig. Imaging*. 21(4), 383.
- Nagy, M., Hanzlicek, P., Zvarova, J., Dostalova, T., Seydlova, M., Hippman, R., Smidl, L., Trmal, J. & Psutka, J. (2008). Voice-controlled Data Entry in Dental Health Record. *Stud. Health Technol. Inform.* 2008;136:529-34
- Oulasvirta, A., Engelbrecht, K-P., Jameson, A., & Möller, S. (2006). The relationship between user errors and perceived usability of a spoken dialogue system. The 2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of Systems. Berlin, Germany.
- Starlander, M., Bouillon, P., Chatzichrisafis, N., Santaholma, M., Rayner, M., Hockey, B.A., Isahara, H., Kanzaki, K. & Nakao, Y. (2005). Practising Controlled Language through a Help System Integrated into the Medical Speech Translation System (MedSLT). In: *Proceedings of the MT Summit X*, Phuket, Thailand.
- Tang, P.C. & McDonald C.J. (2001). Computer-based patient-record systems. In: Shortliffe E.H. & Perreault L.E., eds. *Medical Informatics: Computer Applications in Health Care and Biomedicine*. 2nd ed. New York: Springer, 2001:327-358.
- Teel, M.M., Sokolowski, R., Rosenthal, D. & Belge, M. (1998). Voice-enabled structured medical reporting. In: *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI 98)*, Los Angeles, CA, USA. 595-602.
- Wormek, A.K., Ingenerf, J. & Orthner, H.F. (1997). SAM: Speech-Aware Applications in Medicine to Support Structured Data Entry. In: *Proceedings AMIA Annual Fall Symposium*. 774-778.

Danksagung

Ein Großteil der vorgestellten Arbeiten wurde im österreichischen Kompetenznetzwerk COAST (<http://www.coast.at>) mit Förderung vom Bundesministerium für Wirtschaft, Familie und Jugend sowie vom Zentrum für Innovation und Forschung, Wien, durchgeführt. OFAI erhält Unterstützung von den österreichischen Bundesministerien für Verkehr, Innovation und Technologie sowie für Wissenschaft und Forschung.

Kontaktinformation

Alexandra Klein

E-Mail: vorname.nachname@ofai.at