

Manuel Burghardt, Regensburg

Annotationsergonomie: Design-Empfehlungen für linguistische Annotationswerkzeuge

Dieser Beitrag skizziert die Kluft, die zwischen linguistischen Annotationswerkzeugen und etablierten Methoden des *Usability Engineering* besteht. Eine Evaluationsstudie dreier exemplarischer Annotationstools zeigt unterschiedliche Kategorien von Usability-Problemen auf, welche die Basis für eine Sammlung von 28 Design-Empfehlungen für benutzerfreundliche Annotationswerkzeuge bilden.

Deskriptoren: Benutzerfreundlichkeit, Bewertung, Erschließungsmethode, Annotation

Annotation ergonomics – Design guidelines for linguistic annotation tools

The paper discusses the chasm between the actual usability qualities of linguistic annotation tools and the requirements of current usability engineering standards and practices. A usability study of three selected annotation tools illustrates typical usability problems. A collection of 28 design guidelines for annotation tools is derived from these data.

Keywords: usability, evaluation, indexing, annotation

Ergonomie de l'annotation: recommandations pour les outils d'annotation linguistique

Le présent document décrit le fossé qui existe entre les outils d'annotation linguistiques et les méthodes d'ergonomie informatique éprouvées. Une étude d'évaluation de trois outils d'annotation exemplaires montre différentes catégories de problèmes de facilité d'utilisation qui forment la base d'une collection de 28 recommandations pour la conception d'outils d'annotation conviviales.

Mots-clés: facilité d'utilisation, évaluation, méthode de développement, annotation

1 Linguistische Annotationswerkzeuge und *Humanist-Computer Interaction*

Als grundlegende Möglichkeit implizit vorhandene Informationen explizit verfügbar und damit maschinell verarbeitbar zu machen, spielt Annotation nicht nur im Kontext von *Information Retrieval*-Systemen wie etwa Web-Suchmaschinen (*Semantic Web*) eine wichtige Rolle, sondern zunehmend auch in verschiedenen wissenschaftlichen Disziplinen wie etwa der Literaturwissenschaft (Annotation von Primärtexten) oder der Biologie (Annotation von Genomen). Insbesondere die empirische Sprachwissenschaft, in der sich die Korpuslinguistik als eigene Methode weitestgehend etabliert hat, profitiert von großen Mengen annotierter Daten, welche dann automatisch ausgewertet und visualisiert werden können, um so wichtige Hinweise und Antworten auf linguistische Fragestellungen zu erlangen. Dementsprechend werden bei der linguistischen Annotation Textkorpora zunächst mit sprachlicher Information angereichert, um später komplexe Abfragen an das annotierte Korpus zu ermöglichen (Leech, 1997). Dabei kann ein und derselbe Text parallel auf unterschiedlichen Ebenen annotiert werden: Die Annotation der Wortarten (POS = *parts of speech*) stellt bei den meisten Korpora eine grundlegende Annotations-ebene dar, auf der weitere Ebenen wie etwa syntaktische Annotation oder die Annotation semantischer Phänomene aufbauen. Für einige Bereiche der Linguistik kann die Annotation bis zu einem gewissen Grad durch Lexika und Regeln automatisiert werden, andere linguistische Phänomene müssen hingegen wegen ihrer Komplexität und Kontextabhängigkeit weiterhin von menschlichen Annotatoren markiert werden. Zahlreiche Literaturstellen belegen, dass diese Art der Annotation extrem kostenintensiv und zeitaufwendig ist (vgl. z. B. Dandapat, Biswas, Choudhury, & Bali, 2009; Santos & Frankenberg-Garcia, 2007). Um den mühseligen Prozess manueller Annotation zu unterstützen, gibt es deshalb eine Reihe von linguistischen Annotationswerkzeugen (Burghardt & Wolff, 2009), welche die wesentlichen Arbeitsschritte auf dem Weg hin zu einem annotierten Korpus erleichtern

sollen. Eine grundlegende Systematisierung bestehender Tools kann anhand der jeweils unterstützen Annotationsmodalität erfolgen: So gibt es Tools für die Annotation von geschriebenem Text oder gesprochener Sprache, sowie auch zur Annotation von z. B. Mimik und Gestik bei statischen Bildern oder Videodateien.

Das Hauptziel vieler bestehender Textannotationswerkzeuge ist es technische Details, wie z. B. XML-kodierte Daten, vor dem Annotator zu verbergen. Allein durch diese technische Abstraktion wird allerdings der Annotationsprozess noch nicht wesentlich vereinfacht, vielmehr scheitern technisch unbedarfte, *einfache Annotatoren*¹ an der schlechten Usability der verfügbaren Tools. Die hier zu beobachtende Kluft zwischen Tool-Entwicklern, denen es häufig an Wissen über die eigentliche Anwendungsdomäne und Nutzerschaft ihrer Softwarewerkzeuge fehlt, und Tool-Nutzern, die häufig nicht die technische Expertise zur Bedienung bestehender Tools mitbringen, ist paradigmatisch für viele andere Bereiche der *Digital Humanities*² (Unsworth, 2000). Dabei fällt auf, dass etablierte Methoden des *Usability Engineering*, bei der Entwicklung von Tools und Ressourcen in den Digital Humanities bisher nur wenig Beachtung finden. Dieser Beitrag macht deutlich, dass Usability kein unscharfes Kriterium ist, sondern durch Usability Engineering auch für linguistische Annotationswerkzeuge systematisch umgesetzt werden kann. Durch die Optimierung der Benutzerfreundlichkeit von linguistischen Annotationswerkzeugen soll exemplarisch die Schlüsselfunktion des bislang vernachlässigten Faktors *Usability* für Tools und Ressourcen in den Digital Humanities aufgezeigt werden.

2 Usability-Evaluation bestehender Annotationswerkzeuge

Obwohl die Kosten manueller Annotation sehr hoch sind, wurde der *Usability* von Annotationswerkzeugen bislang nur wenig Beachtung geschenkt (zu den wenigen Ausnahmen gehören etwa Dipper, Götze, & Stede, 2004; Reidsma et al., 2004). Mit Hilfe der Inspektionsmethode des *Heuristic Walkthrough* (Sears, 1997) wurden drei An-

notationswerkzeuge (GATE³, MMAX2⁴, UAM CorpusTool⁵) von jeweils drei Usability-Experten, die zudem grundlegendes Wissen über die Domäne der linguistischen Annotation hatten, evaluiert⁶. Das Ergebnis dieser Evaluation ist eine Sammlung von Usability-Problemen, die von den verschiedenen Evaluatoren für die jeweiligen Tools identifiziert wurden. Nach einem ersten Auswertungsschritt konnten die entdeckten Probleme zu insgesamt 81 unterschiedlichen Usability-Verstößen (*unique problems*) geclustert werden. Diese Probleme lassen sich dabei in zwei grundlegende Kategorien unterteilen: Einerseits allgemeine Usability-Probleme (vgl. z. B. Johnson, 2007), wie sie auch in beliebigen anderen Softwaretools vorkommen können, und andererseits domänen-spezifische Probleme für den Bereich von Annotationstools. Die Gruppe der allgemeinen Usability-Probleme umfasst 30 der insgesamt 81 Probleme (vgl. Abbildung 1). Die identifizierten Probleme beziehen sich dabei oftmals auf fehlendes oder falsches Feedback des Tools an den Benutzer sowie eine irreführende Benutzerführung, die vom System automatisch vorgegeben wird. Eine zweite Problemgruppe stellen unkonventionelle Bedienelemente dar, die es für den Benutzer schwierig machen, schnell und effizient typische Aufgaben zu bewältigen. Die Gruppe der 51 domänen-spezifischen Usability-Probleme von Annotationstools wurde in insgesamt fünf Sub-Kategorien eingeteilt. Die Usability-Evaluation zeigt, dass die unterschiedlichen Tools teilweise höchst unterschiedliche Probleme, und damit individuelle Stärken und Schwächen aufweisen. Eine qualitative Auswertung der Probleme erlaubt schließlich die Formulierung generischer Design-Empfehlungen für Annotationstools, die als Checkliste oder Leitfaden für die Entwickler künftiger Tools herangezogen werden können.

3 Design-Empfehlungen für benutzerfreundliche Annotationstools

Dieser Abschnitt fasst die Ergebnisse der expertenbasierten Evaluationsstudie in Form generischer Design-Empfehlungen für benutzerfreundliche Annotationswerkzeuge zusammen.

³ <http://gate.ac.uk/>

⁴ <http://mmax2.sourceforge.net/>

⁵ <http://www.wagsoft.com/CorpusTool/>

⁶ Für eine detaillierte Beschreibung des Testdesigns sowie der Durchführung siehe Burghardt (2012).

¹ Der Begriff „*plain annotators*“ entstammt einer grundlegenden Benutzerklassifikation von Annotationstoolbenutzern aus Reidsma, Jovanovic, & Hofs (2004)

² Digital Humanities bezeichnen typischerweise geistes- und sozialwissenschaftliche Disziplinen, die bei ihrer wissenschaftlichen Praxis auf digitale Ressourcen oder Tools (Schreibman, Siemens, & Unsworth, 2008) zurückgreifen.

Hauptkategorie	Subkategorie	GATE	MMA2	UAM CT	Gesamt
Allgemeine Usability-Probleme	Feedback und Benutzerführung, Fehlermeldungen	2	6	7	15
	Bedienelemente und Layout	4	3	8	15
Domänen-spezifische Usability-Probleme von Annotationstools	Wording und Metaphern	4	1	2	7
	Primärdaten importieren und editieren	4	2	3	9
	Annotationsschema importieren, editieren oder neu erstellen	7	5	5	17
	Annotationen erstellen, löschen oder editieren	6	3	2	11
	Sekundärdaten visualisieren, insbesondere parallele Annotationen	2	3	2	7
Gesamtzahl der Usability-Probleme pro Tool:		29	23	29	81

Abb. 1: Übersicht über die Gesamtzahl von Usability-Problemen, kategorisiert nach allgemeinen und domänen-spezifischen Problemen sowie zusätzlich nach entsprechenden Sub-Kategorien.

3.1 Wording und Metaphern

Die Empfehlungen in diesem Bereich beschreiben, wie Usability-Probleme, die durch unpassende Benennung oder Metaphern entstehen, vermieden werden können.

E1 Erfinden Sie keine neuen Metaphern für grundlegende Interaktionsparadigmen, die man bereits aus anderen Tools kennt, sondern halten Sie sich an konventionalisierte Benennungen für grundlegende Aktionen wie z. B. das Importieren oder Speichern einer Datei.

E2 Sehen Sie von der Benutzung technischer Ausdrücke ab, auch wenn sie aus der Perspektive eines Entwicklers zunächst plausibel klingen. Versuchen Sie stattdessen technische Konzepte in der Sprache der jeweiligen Domäne zu formulieren.

E3 Stellen Sie sicher, dass die verwendeten Metaphern von der Benutzergruppe aus der Domäne der linguistischen Annotation gut verstanden werden. Falls eine Sammlung von Metaphern eingesetzt wird, sollten diese einerseits konsistent, andererseits jedoch gut voneinander unterscheidbar sein.

E4 Die Hilfe-Funktion sollte ein Wording benutzen, welche das Problem aus der Perspektive des Benutzers beschreibt.

3.2 Primärdaten

Alle Empfehlungen in der Kategorie „Primärdaten“ betreffen Usability-Probleme, die in Zusammenhang mit der Bearbeitung des Originaltexts innerhalb des Annotationstools stehen.

E5 Stellen Sie sicher, dass der Benutzer vom System automatisch durch den Importierungsprozess eines Primärtexts geführt wird, und machen Sie deutlich welche Parameter dabei gesetzt werden müssen. Bieten Sie außerdem Defaultwerte sowie eine Liste mit Auswahlmöglichkeiten (anstelle von Freitextfeldern) für die Konfigurationsparameter an.

E6 Automatisieren Sie die Vorverarbeitungsparameter so weit wie möglich, und stellen Sie sinnvolle Defaultwerte zur Verfügung. Bieten Sie für fortgeschrittene Nutzer die Möglichkeit, die Werte individuell zu konfigurieren.

E7 Stellen Sie eine Vorschau für den importierten Text innerhalb des Tools zur Verfügung. Stellen Sie sicher, dass die Vorschau als solche erkennbar ist und nicht mit dem tatsächlichen Dokument (welches später annotiert werden soll) verwechselt wird.

E8 Stellen Sie sicher, dass der Benutzer optional die Darstellung des Primärtexts innerhalb des Tools an individuelle Präferenzen (Schriftart, Schriftgröße, Farbe, etc.) anpassen kann.

E9 Die Primärdaten sollten klar von den Selektionsmarkern der Annotationsbasen unterscheidbar sein.

E10 Stellen Sie einen Mechanismus zur Verfügung, mit dem man mehrere Dokumente gleichzeitig importieren und innerhalb des Tools organisieren und verwalten kann (Korpusmanagement).

E11 Stellen Sie sicher, dass der Primärtext nicht versehentlich editiert werden kann. Informieren Sie den Benutzer über die möglichen Konsequenzen, die eine Änderung des Primärtexts mit sich bringen kann.

3.3 Annotationsschema

Das Annotationsschema bezeichnet eine Sammlung von Tags, die zur Annotation der Primärdaten verwendet werden dürfen. Ein Annotationstool sollte unter Berücksichtigung der nachfolgenden Empfehlungen grundlegende Möglichkeiten der Erstellung sowie auch der Modifikation von Schemata unterstützen.

E12 Das Tool sollte den Import von existierenden Annotationsschemata erlauben und dem Benutzer gleichzeitig klar machen, welche Anforderungen an das Schema-Format gestellt werden.

E13 Innerhalb des Tools sollte bei Bedarf ein neues Annotationsschema erstellt oder ein bestehendes Schema geändert werden können. Verbergen Sie technische Details, indem Sie einen Schemaeditor mit grafischer Benutzeroberfläche zur Verfügung stellen

E14 Machen Sie deutlich, welches Annotationsschema zu welchem Text gehört, falls es mehrere Annotationsschemata gibt.

E15 Stellen Sie einen kombinierten Mechanismus zur Verfügung, der die Erstellung einer Annotationsebene und der dazugehörigen Tags nicht in zwei separate Arbeitsschritte aufteilt. Daneben sollte es möglich sein, die Annotationsebene als Ganzes zu bearbeiten (Löschen, Umbenennen, Reihenfolge ändern, etc.) sowie auch Änderungen zu einzelnen Tags durchzuführen.

E16 Stellen Sie einen einfachen Mechanismus zur Verfügung, der es erlaubt Tags von einer Annotationsebene auf eine andere zu verschieben.

E17 Da sich das Annotationsschema in vielen Annotationsprojekten ständig weiterentwickelt (Schemaevolution) sollte es eine Möglichkeit geben das Schema bei Bedarf *ad hoc* zu modifizieren.

3.4 Annotation

Die Empfehlungen in diesem Bereich betreffen Usability-Aspekte für den Bereich der tatsächlichen Annotation, d. h. der Selektion der Annotationsbasis, der Auswahl des passenden Tags aus dem Annotationsschema sowie der tatsächlichen Zuweisung, Änderung oder Löschung einer Annotation.

E18 Stellen Sie konventionalisierte Interaktionsmechanismen zur Verfügung, wie sie dem Benutzer bereits von anderen Textverarbeitungsprogrammen bekannt sind.

E19 Bieten Sie eine Option für automatische Segmentierungstools, wie z. B. *Tokenizer* oder *Sentence Splitter*, an. Ermöglichen Sie es dem Benutzer, schnell und un-

kompliziert die automatisch generierten Segmente zu korrigieren.

E20 Das Tool sollte eine einfache Änderung der Reichweite einer Annotationsbasis, oder deren vollständige Löschung erlauben.

E21 Machen Sie das Annotationsschema für eine spezifische Annotationsebene jederzeit verfügbar, um die Auswahl von gültigen Annotationselementen (Tags) zu erleichtern.

E22 Stellen Sie einen einfachen Annotationsmechanismus zur Verfügung, der mit einer minimalen Anzahl von Aktionen (Mausklicks, Tastenanschläge) auskommt: (1) Auswahl einer Annotationsbasis, (2) Auswahl eines passenden Tags aus dem Annotationsschema, (3) Zuweisung der Annotation zur Annotationsbasis.

E23 Stellen Sie einen einfachen Mechanismus zur Auswahl von Tags aus unterschiedlichen Annotationsebenen bereit.

3.5 Sekundärdaten

Die Empfehlungen zur letzten Problemkategorie beziehen sich auf die Sekundärdaten, d. h. die Annotationen zum Primärtext. Probleme entstehen hier vor allem durch eine unpassende bzw. schlecht unterscheidbare Visualisierung von Primär- und Sekundärdaten.

E24 Das Tool sollte Annotationen anzeigen, sobald der Benutzer auf ein annotiertes Element klickt oder mit dem Mauszeiger darüber fährt.

E25 Bieten Sie einen Mechanismus zum Filtern von sichtbaren Annotationen nach einzelnen Tags (z. B. zeige alle Substantive an) sowie nach ganzen Annotationsebenen an (z. B. blende alle POS-Annotationen aus).

E26 Bieten Sie dem Benutzer eine Möglichkeit an, die Darstellung der Annotationen sowie auch der Annotationsbasen individuell zu gestalten.

E27 Stellen Sie eine adäquate Visualisierung für parallele Annotationen zu einer Annotationsbasis bereit (z. B. durch Verwendung einer Stapelmetapher).

E28 Bieten Sie für erfahrene Nutzer eine XML-Ansicht der annotierten Daten an.

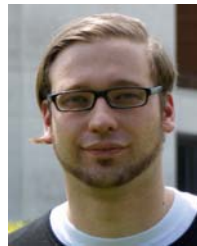
4 Ausblick

Diese Studie zur „Annotationsergonomie“ liefert einen Beitrag zum noch jungen Feld der *Humanist-Computer Interaction*, in dem glücklicherweise der Faktor Usability zunehmend an Bedeutung gewinnt (vgl. z. B. Santos &

Frankenberg-Garcia, 2007). Die Design-Empfehlungen, die aus der Usability-Evaluation dreier exemplarischer Tools abgeleitet werden konnten, mögen als Startpunkt für weiter reichende Studien dienen, deren Endziel eine validierte Sammlung von Usability Patterns für das Design von linguistischen Annotationstools ist⁷. Solch eine Pattern-Sammlung (Borchers, 2001) kann künftigen Tool-Entwicklern dabei helfen, benutzerfreundlichere Werkzeuge zu entwickeln, oder ein *Usability Refactoring* (Garrido, Rossi, & Distant, 2011) bestehender Tools vorzunehmen.

Literatur

- Borchers, J. (2001). A Pattern Approach to Interaction Design. Wiley & Sons.
- Burghardt, M. (2012). Usability Recommendations for Annotation Tools. 6th Linguistic Annotation Workshop – Proceedings of the ACL 2012, 104–112.
- Burghardt, M., & Wolff, C. (2009). Werkzeuge zur Annotation diachroner Korpora. In: W. Hoepfner (Ed.): Proc. GSCL-Symposium Sprachtechnologie und eHumanities, 21–31.
- Dandapat, S., Biswas, P., Choudhury, M., & Bali, K. (2009). Complex Linguistic Annotation – No Easy Way Out! A Case from Bangla and Hindi POS Labeling Tasks. Linguistic Annotation Workshop III, 10–18.
- Dipper, S., Götze, M., & Stede, M. (2004). Simple annotation tools for complex annotation tasks: an evaluation. In: Proceedings of the LREC Workshop on XML-based Richly Annotated Corpora, 54–62.
- Garrido, A., Rossi, G., & Distant, D. (2011). Refactoring for Usability in Web Applications. In: IEEE Software (28), 60–67.
- Johnson, J. (2007). GUI Bloopers 2.0, Second Edition: Common User Interface Design Don'ts and Dos (Interactive Technologies). Morgan Kaufmann.
- Leech, G. (1997). Introducing Corpus Annotation. In R. Garside, G. Leech, & A. McEnery (Eds.), Corpus Annotation. Linguistic Information from Computer Text Corpora, 1–18. Addison Wesley Longman.
- Reidsma, D., Jovanovic, N., & Hofs, D. (2004). Designing Annotation Tools Based in Properties of Annotation Problems. Report for the Centre for Telematics and Information Technology. University of Twente: Dept. of Computer Science, HMI Group, 1–13.
- Santos, D., & Frankenberg-Garcia, A. (2007). The corpus, its users and their needs: a user-oriented evaluation of COMPARA. In: International Journal of Corpus Linguistics, 12(3), 335–374.
- Schreibman, S., Siemans, R., & Unsworth, J. (2008). A Companion to Digital Humanities (Blackwell Companions to Literature and Culture). Wiley-Blackwell.
- Sears, A. (1997). Heuristic Walkthroughs: Finding the Problems Without the Noise. In: International Journal of Human-Computer Interaction 9(3), 213–234.
- Unsworth, J. (2000). Scholarly Primitives: what methods do humanities researchers have in common, and how might our tools reflect this? Symposium on “Humanities Computing: formal methods, experimental practice.” <http://www3.isrl.illinois.edu/~unsworth/Kings.5-00/primitives.html> [23.7.2012].



Manuel Burghardt, M. A.
Lehrstuhl für Medieninformatik
Universität Regensburg
Universitätsstraße 31
93053 Regensburg
manuel.burghardt@ur.de

Manuel Burghardt studierte Informationswissenschaft und Linguistik an der Universität Regensburg. Derzeit ist er dort als wissenschaftlicher Mitarbeiter am Lehrstuhl für Medieninformatik tätig. Er befasst sich in Forschung und Lehre mit den Themen Digital Humanities, Korpuslinguistik und Mensch-Maschine-Interaktion.

⁷ Begleitend zur hier vorgestellten Evaluationsstudie entsteht derzeit eine Dissertation zum Thema „*Engineering Annotation Usability*“ an der Universität Regensburg.