

# **InterpretBank: Ein Tool zum Wissens- und Terminologiemanagement für Simultandolmetscher**

*Claudio Fantinuoli*

*Johannes Gutenberg-Universität Mainz, Gernersheim, Deutschland*

`fantinuo@uni-mainz.de`

## **1. Einleitung**

### **1.1 Nutzergruppen und Aufgabenbereiche**

Erfolgreiches Dolmetschen setzt qualifizierte Vorbereitung voraus. Dazu gehört auch die nutzeradäquate Gestaltung und kontinuierliche Pflege von Terminologiebeständen sowie die Möglichkeit auf Informationen und Terminologie schnell und effizient zugreifen zu können. Als Fachgebietslaien müssen sich professionelle Dolmetscher auf einen bevorstehenden technischen Einsatz gezielt vorbereiten und sowohl fachliches als auch terminologisches Wissen aneignen. Da aufgrund der Fachlichkeit der Themen keine adäquaten Ressourcen zur Verfügung stehen, muss der Dolmetscher sie selbst erarbeiten.

Um diese Prozesse zu optimieren wird basierend auf den neuesten Erkenntnissen der Dolmetschwissenschaft und der Computerlinguistik der Versuch unternommen, eine speziell für die Dolmetscher konzipierte Anwendung Namens *InterpretBank* zu entwickeln und zu implementieren.

In diesem Artikel sollen die relevanten Grundlagen der Dolmetschwissenschaft und der Computerlinguistik sowie zwei der Module von *InterpretBank* vorgestellt werden.

### **1.2 Elektronische Datenverarbeitung**

Während terminologische Daten und fachliche Informationen lange Zeit auf Papier verfasst und verbreitet wurden, bieten computerlinguistische Anwendungen und das Internet neue Möglichkeiten der Datenverarbeitung und -darstellung. Die Verfügbarkeit großer Mengen an frei zugänglichen Fachtexten im Internet (Kaligarriff und Grafenstette 2003), die dynamische Datendarstellung und die unterschiedlichsten Möglichkeiten des Datenzugriffs mittels ausgereifter Suchverfahren sind nur einige der wichtigsten Vorzüge der elektronischen Datenverarbeitung.

Die starre und meist normative Struktur der gedruckten lexikographischen Werke, wie z.B. Wörterbücher und Lexika, oder die Unflexibilität der gedruckten Fachtexte, z.B. Fachzeitschriften oder -bücher, überlassen den dynamischen und

linguistisch deskriptiven Ansätzen der computerunterstützten Wissens- und Terminologieverwaltung das Feld. Die Vernetzung von kontrollierten Datenbeständen (Glossare) mit automatisch gesammelten Fachtexten (Korpora), die automatische Extraktion von relevanten Fachbegriffen (Fantinuoli 2006), die Einbindung von Datensammlungen in speziell für die Bedürfnisse der Anwender programmierten Anwendungen, um z.B. aus den gewonnenen Datenmengen gezielt Informationen terminologischer und inhaltlicher Natur abrufen zu können, sowie die dynamische Erzeugung von Darstellungsformen (z.B. in der Form von semantischen Netzen oder Mind Maps), können die Möglichkeiten der *Knowledge Experience* - der Anneigung von Wissen und Terminologie - erweitern und ergänzen.

### 1.3 Computereinsatz bei Dolmetschern

In den letzten Jahren wurden zahlreiche Umfragen unter Konferenzdolmetschern zu deren Erfahrung mit Computern und Terminologietools durchgeführt. Die Arbeiten von Will (2000) und Valentini (2002) und die Umfrage vom Sprachen & Dolmetscher Institut München (2007) zeigen ein ernüchterndes Bild hinsichtlich der Verbreitung der für Dolmetscher entwickelten Tools. Die meisten Befragten benutzen heutzutage für die Terminologieverwaltung immer noch selbstproduzierte Lösungen, wie z.B. Word- oder Excel-Tabellen, nur selten aber Tools, die auch eine Simultanmodalität besitzen, so gut wie niemals computerlinguistische Anwendungen wie zum Beispiel Concordancer oder Terminologieextraktionswerkzeuge. Gleichzeitig aber zeichnet sich im Laufe der Jahre ein immer wachsendes Interesse für Computeranwendungen ab, die die Vorbereitung, Durchführung und Nachbearbeitung eines Dolmetscheinsatzes effizienter gestalten sollten. Nach dem gleichem Modell der Dolmetschwissenschaft, die zwischen *advance*, *last-minute* und *in-conference preparation* (Gile 1995) unterscheidet, muss eine dolmetschergerechte Anwendung die unterschiedlichen Phasen eines Dolmetscheinsatzes von der Vorbereitung bis zum Kabineneinsatz und zur Nachbearbeitung berücksichtigen und zum Tragen bringen.

### 1.4 Modelle und Methoden der Dolmetschwissenschaft

Die Dolmetschwissenschaft ist sich einig, dass

[ ] *before starting a conference, interpreters should acquire as much specific knowledge as possible (Gile 1995)*

Die Vorbereitungsphase einer Fachkonferenz in einem dem Dolmetscher noch nicht bekannten Fachgebiet spielt eine entscheidende Rolle im Leben eines professionellen Dolmetschers. In dieser Phase muss sich der Dolmetscher eine Reihe von Informationen sprachlicher und inhaltlicher Natur aneignen, die notwendig sind um einen Dolmetscheinsatz erfolgreich durchzuführen.

Die terminologische und fachliche Vorbereitung wird von Kalina (1998) als Strategie beschrieben, die die kognitiven Prozesse auf die Zeit vor der Konferenz vorverlegt und somit den Dolmetscher während der Verdolmetschung entlastet.

Die gewonnenen freien Kapazitäten bedeuten in der Regel ein besseres Output der Verdolmetschung. Darüber hinaus spielt die Anwendung einer nutzeradäquate Terminologie in sehr fachlichen Konferenzen eine weitere wichtige Rolle: Einerseits schafft sie Vertrauen dem Dolmetscher gegenüber, andererseits vereinfacht sie die Fachkommunikation und erlaubt dem Dolmetscher präzisere und knappere Formulierungen. Und das verbessert wiederum die Qualität bzw. die Wahrnehmung der Qualität der Verdolmetschung.

Um sich in kürzester Zeit auf neue und komplexe Themen vorbereiten zu können, bedient sich der Dolmetscher unterschiedlicher Techniken. Im wesentlichen beschränken sie sich aber auf die Bearbeitung der zur Verfügung gestellten Unterlagen und auf die Internetrecherchen. Diese Operationen werden vorwiegend sehr traditionell durchgeführt und zwar auf dem Medium Papier (Valentini 2002).

Nachdem die Dolmetschwissenschaft sich zuerst auf die reine Beschreibung des Verhältnisses zwischen Dolmetschern und Computern, bzw. Vorbereitungs-mustern beschränkt hat, erleben wir erst in den letzten Jahren, dass die Dolmetschwissenschaft einen Wandel vom rein deskriptiven zu einem eher konstruktiven Ansatz durchgemacht hat. Rütten (2007) zum Beispiel befasst sich auf eine detaillierten Art und Weise mit den strukturellen Prozessen der Organisation terminologischer Arbeit für Dolmetscher, während Will (2007) den Versuch unternimmt, Modelle und Methoden dafür zu entwickeln. Diese theoretischen dolmetschbezogenen Ansätze gekoppelt mit den Erkenntnissen der Korpuslinguistik im Bereich Spracherwerb und Übersetzungswissenschaft sowie der Computerlinguistik dienen als Grundlage für die Implementierung von *InterpretBank*.

## 2. InterpretBank - Die Struktur

InterpretBank ist ein neues, offenes und modulares Tool, welches den Dolmetscher im Bereich Wissens- und Terminologiemanagement vor, während und nach der Konferenz unterstützt. Die einzelnen Bauteile - die Module - sind unabhängige Anwendungen, die eine bestimmte Aufgabe erfüllen.

Unter anderem sind folgende Module vorgesehen:

- *ConferenceMode*: Modul zum Nachschlagen von Glossaren während des Simultaneinsatzes
- *TermMode*: Modul zur Erstellung und Pflege der Terminologiebestände
- *CorpusMode*: Modul zur automatischen Termextraktion und Informationssuche aus automatisch hergestellten Fachkorpora
- *DrillMode*: Modul zum aktiven und dolmetschspezifischen Lernen der Glossare

Alle Module sind mittels offener Formate miteinander und ggf. mit externen Anwendungen verbunden. Die im *TermMode* verwalteten terminologischen

Einträge sind z.B. mit dem *Concordancer* des *CorpusMode* verbunden und ermöglichen somit das direkte Starten einer Kontextsuche aus einem bestimmten Glossar; aus dem *ConferenceMode* kann man neue Termini, die man während der Verdolmetschung zu den bestehenden Glossaren hinzufügen will, direkt eingeben ohne die Konferenzmodalität verlassen zu müssen, usw. Um eine möglichst hohe Akzeptanz seitens der Dolmetscher zu erzielen, muss *InterpretBank* hohe Flexibilität und Leistungsfähigkeit mit einer einfachen, benutzerfreundlichen und selbsterklärenden Bedienungsoberfläche in Einklang bringen (Valentini 2002). Dieser kritischer Balanceakt stellt für Design und Implementierung eine große Herausforderung dar.

Im Folgenden werden die wichtigsten Funktionalitäten zweier Module, *ConferenceMode* und *CorpusMode*, näher erläutert, die aus dolmetschwissenschaftlicher und computerlinguistischer Sicht innovativer sind.

## 2.1 ConferenceMode

*ConferenceMode* ermöglicht dem Konferenzdolmetscher den schnellen und bedarfsorientierten Zugriff auf bestehende mehrsprachige Terminologiedateien. Bedingt durch die Eigenschaften des Dolmetschprozesses - Zeitdruck, hohe Konzentration, etc - muss die Anwendung für den Einsatz in der Kabine vor allem Wert auf folgende Grundbeschaffenheiten legen (SDI 2007):

- schnelle und flexible Suchfunktion
- Übersichtlichkeit
- komfortable und schnelle Eingabe neuer Termini
- intuitive Bedienbarkeit
- Kompatibilität mit anderen Programmen

*ConferenceMode* verwendet als Datenbank eine einzige Txt-Datei, das so genannte *Aktive Glossar*. Diese Datei enthält alle Wortpaare, die im Vorfeld für einen Einsatz geladen wurden und bleibt unverändert bis *ConferenceMode* für den nächsten Einsatz mit einem neuen Glossar geladen wird. Diese Lösung ermöglicht es dem Dolmetscher das aktive Glossar zusammenzustellen, indem er ein oder mehrere Glossare aus *TermMode* oder aus anderen Programmen (MS Word, MS Excel, Multiterm, CSV-Dateien, etc.) nacheinander lädt. Dank dieser hohen Flexibilität kann der Dolmetscher am Einsatzort schnell und unproblematisch Glossare von Kunden oder Kollegen lesen und zum aktiven Glossar hinzufügen.

Auch die Eingabe von neuen Termini während des Einsatzes erfolgt schnell und komfortabel durch eine eigene Maske. Die neuen Termini werden direkt dem aktiven Glossar hinzugefügt, sodass diese gleich abrufbar sind. Damit diese Termini auch im eigenen Terminologieverwaltungstool eingetragen werden, werden sie in einer separaten Datei oder direkt in *TermMode* aufgenommen.

Um den Dolmetschprozess so wenig wie möglich zu beeinträchtigen und den Dolmetscher bei der Suche nach passenden Fachbegriffen auch während der Verdolmetschung zu unterstützen, ist es notwendig, dass einerseits das erforderliche Input so klein wie möglich ist und dass andererseits das Output so übersichtlich wie möglich gestaltet und auf das Minimum reduziert wird. Ziel ist es, dass der Dolmetscher mit wenig Aufwand möglichst wenige aber gleichzeitig präzise Treffer angezeigt bekommt. Der gesuchte Begriff wird mittels Tastatur eingegeben, die Suche mit der Entertaste oder mit dem eigenen Suchalgorithmus (ohne Entertaste) begonnen. Dieser ermöglicht das Anzeigen der Treffer schon während der Eingabe. Sobald die eingestellte Anzahl von Treffern angezeigt wird, wird die Eingabemaske für die nächste Suche freigegeben. Dank dieser interaktiven Suchmethode wird der Dolmetscher bei der Suche erheblich entlastet. Die Reduzierung der angezeigten Treffer erfolgt u.a. durch den Einsatz von *Stopwords*. Wenn man zum Beispiel nach dem Wort "Dermatologie" sucht und die Buchstabenkette "d", "de" oder "der" eingibt, wird der Eintrag "Entzündung der Bauchspeicheldrüse" nicht angezeigt, weil der Artikel "der" auf der Stopwortliste steht.

## 2.2 CorpusMode

*CorpusMode* unterstützt den Dolmetscher während der Vorbereitungsphase durch die Bereitstellung von fachlicher Terminologie und durch die Möglichkeit in fachbezogenen Textsammlungen zu stöbern, um inhaltliche und terminologische Fragen zu vertiefen.

Das Grundprinzip von *CorpusMode* ist einfach: Aus wenigen Fachbegriffen - zum Beispiel aus Titeln oder Abstracts der Konferenzteilnehmern - wird automatisch nach der Methode von BootCaT (Baroni und Bernardini 2004) eine Sammlung von Fachtexten aus dem Internet heruntergeladen und als einfache Text-Dateien aufbereitet. Das erstellte Korpus dient dann zur automatischen Extraktion der Fachterminologie, die auf der Grundlage von statistischen und linguistischen Kriterien durchgeführt wird.

Ausgehend von den extrahierten Termkandidaten oder von den Fachtermini, die im Laufe der Einsatzvorbereitung getroffen werden, kann der Dolmetscher eine Suche im Fachkorpus starten. Die Ergebnisse dieser "browsing activity", (Bernardini, 2000 / Johns, 1988), die in Form einer Auflistung von sogenannten *Key Words In Context* (KWIK) angezeigt werden, können Informationen über Terminologie, Inhalt und Gebrauch von Begriffen in einem bestimmten Kontext verraten (Zanettin, 1998).

Sammlungen von Fachbegriffen, die das Ergebnis einer automatischen Termextraktion darstellen, können zum Beispiel als Basis für eine gezielte Vorbereitung verwendet werden:

*The process of "knowledge acquisition/language learning" needed by interpreters in order to prepare themselves for a conference can be optimized by making it "terminology-driven", or "bottom-up": from the terminology to the conceptual structure of a*

*particular domain. (Fantinuoli 2006)*

Korpora können als Grundlage für einen nahezu unendlichen "serendipity process" (Johns 1988) dienen, da ein Wort oder ein Satz zum nächsten führen kann. Durch diesen Ansatz kann der Dolmetscher das Korpus - ausgehend von den Fachbegriffen - dynamisch erkunden und dabei Begriffe und deren Bedeutung lernen.

Für eine empirische Untersuchung der Qualität der automatisch erstellten Korpora und der extrahierten Terminologie im Bezug auf die Dolmetscherbedürfnisse wird hier auf die Arbeit von Fantinuoli (2006) verwiesen.

Nach dem Prinzip der Modularisierung verfügt *CorpusMode* über eine Rückkopplung zu den anderen Modulen. So ist es z.B. möglich, aus einem Eintrag in *TermMode*, dem Terminologie-Verwaltungsmodul von *InterpretBank*, direkt eine KWIK-Suche zu starten und eine Serie von Beispielen für das betroffene Wort zu erhalten. Das erspart einerseits viel Arbeit bei der Eingabe der Einträge, da Felder wie Definition oder Beispielsätze in den meisten Fällen nicht mehr nötig sind, andererseits erweitert es die Informationen, die mit einem Eintrag verbunden sind.

## Ausblick

Während Computeranwendungen zum festen Bestandteil des Übersetzerberufs geworden und mittlerweile nicht mehr wegzudenken sind, bleibt das Dolmetschen von den neuesten Entwicklungen und Erkenntnissen im Bereich Translationswissenschaft und Computerlinguistik weiterhin unberührt. Da die möglichen Vorteile eines computergestützten Ansatzes auch für Dolmetscher auf der Hand liegen, will der Autor mit *InterpretBank* versuchen, eine erste Brücke zwischen diesen Disziplinen zu schlagen und praktizierenden oder angehenden Dolmetschern eine Lösung zu bieten, um die Qualität der eigenen Dienstleistung zu steigern und die kognitiven Kompetenzen zu fördern.

## Bibliographische Angaben

- Baroni, M. / Bernardini, S. (2004). BootCaT: Bootstrapping Corpora and Terms from the Web, in: *Proceedings of LREC*, Lisbon: ELDA, S. 1313-1316.
- Bernardini, S. (2000). *Competence, capacity, corpora*. Bologna: CLUEB.
- Gile, D. (1995). *Basic concepts and models for translator and interpreter training*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Kalina, S. (1998). Kognitive Verarbeitungsprozesse, in: Snell-Hornby, M. / Hönig, H. G. / Kußmaul, P. / Schmitt, P. A. (eds.) *Handbuch Translation*. Tübingen: Stauffenburg, S. 330-335.
- Kilgariff, A. / Grefenstette, G. (2003). Introduction to the web as corpus, in: *Computational Linguistics*, 29. S. 333-347.

- Fantinuoli, C. (2006). Specialized Corpora from the Web for Simultaneous Interpreters, in: Baroni, M. / Bernardini, S. (eds.). *Wacky! Working papers on the Web as Corpus*. Bologna: GEDIT. S. 173-190.
- Johns, T. (1988). Whence and whither classroom concordancing? In: Bongaerts et al. (eds.). *Computer applications in language learning*. Foris. S. 9–32.
- Rütten, Anja (2007). *Informations- und Wissensmanagement im Konferenzdolmetschen*. Frankfurt: Lang.
- Valentini, C. (2002). *Uso del Computer in Cabina di Interpretazione*. Tesi di Laurea, SSLiMIT, Bologna.
- Will, M. (2000). Bemerkungen zum Computereinsatz beim Simultandolmetschen. In: Kalina, S. / Buhl, S. / Gerzymisch-Arbogast, H. (eds.). *Dolmetschen: Theorie-Praxis-Didaktik*. St. Ingbert: Röhrig. S. 125-135.
- Will, M. (2007). Terminology Work for Simultaneous Interpreters in LSP Conferences: Model and Method. *MuTra 2007 – LSP Translation Scenarios: Conference Proceedings* Sprachen & Dolmetscher Institut München (2007). Terminologietools für den Einsatz in der Simultankabine. In: *MDÜ 3/2007*. SS. 26 ff.