

Inkrementelle und erwartungsgesteuerte Verarbeitung beim Maschinellen Dolmetschen*

Alexandra Klein, Susanne J. Jekat, Jan W. Amtrup
Arbeitsbereich Natürlichsprachliche Systeme
Fachbereich Informatik
Universität Hamburg
Vogt-Kölln-Str. 30
22527 Hamburg

Das menschliche Dolmetschen muß als kognitive Leistung betrachtet werden, die auf dem komplexen Zusammenspiel verschiedenster Prozesse auf unterschiedlichen perzeptiven und kognitiven Ebenen beruht (Wadensjö, 1992). Im Gegensatz zu der vielschichtigen Interaktion bei der Verarbeitung steht die Strukturierung der Informationen: Eingabe ist ein gesprochener Quelltext, der unter Umständen verschiedene Performanzphänomene enthalten kann, und als Ausgabe wird, abhängig von den Rahmenbedingungen, ein möglichst wohlgeformter gesprochener Text in der Zielsprache erzeugt.

Herkömmliche Ansätze zum maschinellen Dolmetschen nehmen die intuitiv leicht zugängliche Abbildung der quellsprachlichen Äußerung auf die zielsprachliche Äußerung zum Ausgangspunkt und versuchen, diese Abbildung durch das sequentielle Durchlaufen der traditionellen linguistischen Analyseebenen zu modellieren. Der Prozeß des maschinellen Dolmetschens wird dabei analog zum Verhältnis zwischen quell- und zielsprachlicher Struktur als symmetrische Beziehung zwischen einem Sprachverstehens- und einem Sprachproduktionsprozeß behandelt. Maschinelle Dolmetschsysteme erheben daher keineswegs den Anspruch, die menschliche Dolmetschleistung vollständig maschinell zu simulieren, um sie gegebenenfalls gar zu ersetzen.

Die vergrößerte symmetrische Systemkonzeption bereitet Probleme bei der Verarbeitung, wenn inkrementelle Mechanismen integriert werden sollen. Zwar gibt es interessante Ansätze zur inkrementellen Sprachproduktion bzw. -generierung (Abb et al., 1993; Kilger and Finkler, 1994), doch Erkenntnisse hieraus werden in der maschinellen Übersetzung bisher kaum angewendet, was in dem erwähnten Symmetriekonzept von Analyse/Verstehen und Generierung/Produktion begründet liegt. Naturgemäß ist es in einer solchen Systemarchitektur schwierig, die Analyseschritte und quellsprachlichen Einheiten zu jedem Zeitpunkt mit den Generierungsschritten und zielsprachlichen Übersetzungen inkrementell zu koordinieren.

Es ist dagegen unser Ziel, in einem sehr eingeschränkten Bereich Teilprozesse zu modellieren und zu untersuchen, die die menschliche Dolmetschleistung verbessern. Da menschliche Dolmetschprozesse wenig erforscht sind und sich dem unmittelbaren Zugang entziehen, ist allerdings zu erwarten, daß Hypothesen über den menschlichen Dolmetschprozeß, die auf diese Weise in die Modellierung einfließen, Aufschlüsse über letzteren liefern.

*This research was funded by the Federal Ministry of Education, Science, Research and Technology (BMBF) in the framework of the VERBMOBIL Project under Grant 01 IV 101 A/O

Im folgenden soll ein Ansatz vorgestellt werden, der für die sehr kleine, dabei aber gut dokumentierte Domäne Terminabsprache (Projekt Verbmobil) versucht, einige Verarbeitungsprinzipien des Humandolmetschens in einem System zum maschinellen Dolmetschen zu modellieren. Die Verarbeitung erfolgt dabei nicht sequentiell und strikt modular, sondern inkrementell und interaktiv. Auf das Symmetrie-Konzept greifen wir im Bereich der Koordination der quell- und zielsprachlichen Verarbeitungseinheiten zurück. Diese Konzeption wurde im Zusammenhang mit Forschungsergebnissen aus der Translationswissenschaft und der Beschäftigung mit Verarbeitungsmechanismen und Architekturoptionen entwickelt. Teile des Ansatzes sind bereits Bestandteil eines Transfer-Systems in der Domäne Terminabsprache.

Als zentrale Leistungen beim menschlichen wie beim maschinellen Dolmetschen ist die

- Segmentierung des Quelltextes in Übersetzungseinheiten und die
- Koordination der verschiedenen perzeptiven und kognitiven Prozesse

zu betrachten (Amtrup and Jekat, 1995). Diese beiden Aufgaben können nur vor dem Hintergrund einer stark erwartungsgetriebenen Verarbeitung bewältigt werden. Die Einbeziehung von Erwartungen wird ermöglicht durch das strikt inkrementelle Paradigma, das für die menschliche Informationsverarbeitung zumindest im sprachlichen Bereich angenommen wird; nur, wenn verschiedene Bestandteile eines prozessualen Systems an heterogenen Ebenen des Wissens arbeiten, sind Interaktionsphänomene denkbar, die zu einer größeren Effektivität und Robustheit des Gesamtsystems führen. Zudem sind einige Aspekte menschlicher Sprachperformanz, wie die Unterbrechung eines Partners im Dialog oder das Simultandolmetschen, nur mit Hilfe eines inkrementellen Mechanismus möglich. Aus einer anderen Sichtweise, die nicht primär kognitionswissenschaftlich orientiert ist, sondern aus der Ingenieurstradition herrührt, ergeben sich durch die Einführung von Inkrementalität in Programme zur Informationsverarbeitung bedeutende systemtechnische Vorteile durch die Möglichkeit, intermodulare Parallelität zu nutzen.

Es ist zu untersuchen, wie die wichtigen Faktoren Segmentierung, Informationskoordination und die Integration von Erwartungen an den Dialogverlauf in ein System einfließen können und wie ihr Zusammenspiel simuliert werden kann.

Die erwartungsgetriebene Verarbeitung soll auf der Grundlage von Dialogakten (Jekat et al., 1995) simuliert werden. Dialogakte repräsentieren den propositionalen Gehalt einer Äußerung, wie er im Zusammenhang einer bestimmten Gesprächssituation repräsentiert werden kann. Auf die Sachverhaltsrepräsentationen (z.B. Kalender) wird nur zurückgegriffen, wenn eine Übersetzung allein aufgrund der Dialogrepräsentation nicht möglich ist. Die Verarbeitung wird vereinfacht, da Wissensquellen nach dem Prinzip der Variablen Analysetiefe (Prah, Petzold, and Heizmann, 1995) nur bei Bedarf konsultiert werden müssen. Als Entscheidungskriterium bei der Koordination der Informationen dient hierbei die Frage, ob eine Übersetzung zu dem gegebenen Zeitpunkt bereits möglich ist oder ob noch zusätzliche Informationen benötigt werden, die zuerst angefordert werden müssen. In diesem Kontext kann untersucht werden, wie sich bei der Verarbeitung geprochener Sprache die Repräsentation von Sachverhalten zur Repräsentation von Textpropositionen auf sprachlicher Ebene verhält und in welchen Fällen oder Situationen welcher Typ von Wissen benötigt wird.

Vor dem Hintergrund der erwartungsgetriebenen Verarbeitung kommt der inhaltsbezogenen Segmentierung von Gesprächseinheiten für das Dolmetschen besondere Bedeutung zu. Wesentlich sind hier drei Fragestellungen:

(1) Wie kann die Äußerung sinnvoll in Teile aufgeteilt werden, die sich als Ausgangspunkte für Analyse/Verstehen und Generierung/Produktion eignen? (2) Wie wird widersprüchliche Information in einem Äußerungsteil behandelt? (3) Wie kann die in einer Äußerung gegebene Information hierarchisch gegliedert und repräsentiert werden?

Die einer Äußerung zugeordneten Dialogakte bilden die Informationsstruktur einer Gesprächseinheit ab und richten sich nach den Hinweisen, die aus den Wissensquellen auf den unterschiedlichen Analyseebenen (wie zum Beispiel Prosodie, Schlüsselwörter) gewonnen werden können. Gleichzeitig stehen Dialogakte für bestimmte Gesprächsmuster, für die es auf den unterschiedlichen sprachlichen Ebenen Befunde gibt. Besonders aufschlußreich sind in diesem Zusammenhang Merkmale der gesprochenen Sprache, die bis zum gegenwärtigen Zeitpunkt in der Sprachverarbeitung oft vernachlässigt werden, z.B. Routineformeln, Diskurspartikel, Häsitationen etc.. In diesem Zusammenhang wird auch deutlich, daß die wörtliche Übersetzung nicht als adäquate Strategie angesehen werden kann, da Performanzphänomene im spontansprachlichen Quelltext zu redundanten Elementen im Zieltext führen würden.

Die vorgestellte zyklische Approximation an einen einheitlichen Befund für das Äußerungsprofil wird durch eine enge Kopplung von Dialogaktrepräsentation und sprachlichen Wissensquellen simuliert. Bei der inkrementellen Verarbeitung von Gesprächseinheiten werden ständig die in Frage kommenden Wissensquellen aktiviert und mit der bisher aufgebauten Repräsentation der Proposition im Zusammenhang eines Dialogakts verglichen. Ein Informationstransfer im kompletten System gilt dann als vollständig und korrekt, wenn aus der Dialogaktrepräsentation eine der quellsprachlichen Äußerung funktional adäquate zielsprachliche Übersetzung in geprochener Sprache generiert werden kann, die im Einklang mit den Befunden steht, die die verschiedenen Wissensquellen zu den unterschiedlichen Zeitpunkten geliefert haben. Entsprechend den strukturellen und konventionellen Anforderungen der Zielsprache wird dabei gegebenenfalls der quellsprachliche Text erweitert (achievement) oder reduziert (reduction). Hiermit ist gleichzeitig die Modellierung zweier wichtiger menschlicher Dolmetschstrategien (Prahl, Petzold, and Heizmann, 1995) gegeben.

Ein System, das mit Hilfe verschiedener integrierter Wissensquellen Dialogaktrepräsentationen aufbaut und inkrementell verarbeitet, wird Aufschluß über Verarbeitungsmechanismen des menschlichen Dolmetschens geben und zeigen, inwieweit diese innerhalb eines maschinellen Dolmetschsystems eingesetzt werden können. Dabei unterstützt die Relation zwischen quell- und zielsprachlicher Äußerung vor dem Hintergrund des Dialogaktkonzepts die Einsicht in die Interaktion von perzeptiven und kognitiven Prozessen.

Literatur

- Abb, B., Carsten Günther, Michael Herweg, Kai Lebeth, Claudia Maienborn, and Andrea Schopp. 1993. Incremental Syntactic and Phonological Encoding: An Outline of the SYNPHONICS Formulator. In *Proc. of the 4th European Language Generation Workshop*, pages 19–30, Pisa.
- Amtrup, Jan W. and Susanne J. Jekat. 1995. Segmentation of Spoken Language for NLP. In *Proceedings of the KI95 Workshops*, pages 298–299, Bielefeld, September.
- Jekat, Susanne J., Alexandra Klein, Elisabeth Maier, Ilona Maleck, Marion Mast, and J. Joachim Quantz. 1995. Dialogue Acts in VERBMOBIL. Verbmobil Report 65, Universität Hamburg, April.
- Kilger, Anne and Wolfgang Finkler. 1994. Tag-based incremental generation. Draft, DFKI, Saarbrücken.

Prahl, Birte, Susanne Petzold, and Susanne Heizmann. 1995. Variable Analysetiefe und Bewertungskriterien in VERBMOBIL: Translationswissenschaftliche Grundlagen. Verbmobil Report 54, Universität Hildesheim, January.

Wadensjö, C. 1992. Interpreting as interaction. Linköping studies in arts and science, Linköping University.