SEMANTISCHE INTERPRETATION UND ARTIKULATION MIT ÄUSSERUNGS-BEDEUTUNGS-TRANSDUKTOREN

Jens Lindemann

Brandenburgische Technische Universität Cottbus - Senftenberg
lindemann@b-tu.de

Kurzfassung: Auf Basis des in [8] beschriebenen bedeutungsorientierten Sprachmodells erfolgt die semantische Interpretation einer Sprachäußerung in unserer Systemvorstellung durch sogenannte *Äußerungs-Bedeutungs-Transduktoren*. Diese ordnen einem syntaktischen Ausdruck mögliche Bedeutungen zu. Als Bedeutungsrepräsentationssprache dienen *Merkmal-Werte-Relationen*. Anhand einer Beispielanwendung wird in diesem Beitrag neben der Modellierung von *Äußerungs-Bedeutungs-Transduktoren* u. a. auch die Verwendung dieser semantischen Grammatiken für die Sprachgenerierung untersucht. Dazu wird die interpretierte Bedeutungsstruktur mittels des *inversen Äußerungs-Bedeutungs-Transduktors* wieder in mögliche Wortfolgen der Zielsprache übersetzt. Die semantische Sprachverarbeitung erfolgte hier sowohl für die deutsche Lautsprache als auch für die Deutsche Gebärdensprache. Für die Automatenoperationen wurde die Bibliothek *OpenFst* genutzt.

1 Einleitung

Sprache ist die bedeutendste und natürlichste Kommunikationsform. Sprachverarbeitende Systeme sollten in der Lage sein, sinnvolle artikulierte Sprachäußerungen zu „verstehen“. Das Ziel der semantischen Interpretation ist die automatische Generierung einer vom Kontext der Sprechsituation unabhängigen Bedeutungsrepräsentationen zu der analysierten Sprachzeichenfolge. In unserem Systemmodell wird die zugeordnete Bedeutung als *Merkmal-Werte-Relation* (MWR) [7] repräsentiert. *Merkmal-Werte-Relationen* ermöglichen eine hierarchische Strukturierung von semantischen Konzepten und können durch gerichtete azyklische Graphen dargestellt werden.
In der übergeordneten Ebene des Sprachdialogsystems erfolgt eine Weiterverarbeitung der generierten Bedeutungshypothesen. Damit sind z. B. eine Einschränkung der möglichen Bedeutungen sowie die Aktualisierung des Informationsstatus möglich. Anstelle dieser weiteren Verarbeitungsschritte und der daraus resultierenden Inhaltsfestlegung für die Systemreaktion wird in dem vorgestellten Beispielsystem eine Bedeutungsrepräsentation direkt wieder artikuliert. Dieser Artikulationsprozess wird auch als *Text- bzw. Sprachgenerierung* bezeichnet. Dabei werden der ausgewählten Informationsstruktur mögliche Sprachzeichenfolgen zugeordnet. Die Abbildung 1 zeigt die beteiligten semantischen Verarbeitungsschritte der in diesem Beitrag vorgestellten Anwendung. Im diesem Fall kann die Aufgabe der Artikulation als Rückübersetzungsprozess angesehen werden. Folglich lassen sich die für die semantische Interpretation verwendeten *Äußerungs-Bedeutungs-Transduktoren* T_{UM} auch für die Artikulation verwenden. Durch Vertauschung der Ein- und Ausgabezeichen an den Zustandsübergängen entsteht ein *inverser Äußerungs-Bedeutungs-Transduktor* T_{UM}^{-1}. Dieser generiert zu den akzeptierten Konzeptsequenzen die dazugehörigen Wortfolgen der Zielsprache.
Auf diese Weise erhält man als Ergebnis der beiden Übersetzungsprozesse eine sinngemäße

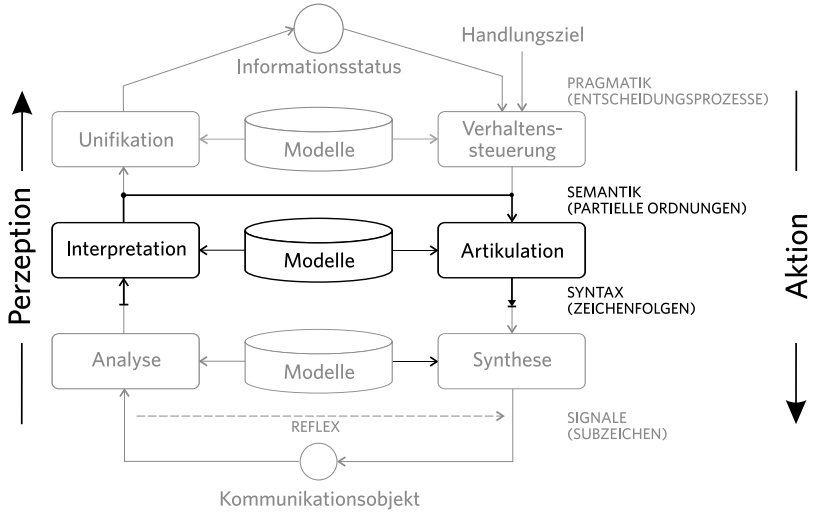

Abbildung 1 - Interpretation und Artikulation mittels Äußerungs-Bedeutungs-Transduktoren in intelligenten hierarchischen Systemen

Wiedergabe (Paraphrase) der analysierten Sprachäußerung. Die Interpretation und Artikulation wurde sowohl für die Deutsche Lautsprache als auch für die Deutsche Gebärdensprache untersucht. Wenn unterschiedliche Quell- und Zielsprachen in den Systemeinstellungen festgelegt sind, dann erreicht man auf diesem Weg eine bedeutungsgetriebene Übersetzung eines erkannten sprachlichen Ausdrucks.

2 Modellierung von Äußerungs-Bedeutungs-Transduktoren

Für die Beispielanwendung wurden *Äußerungs-Bedeutungs-Transduktoren* für die deutsche Lautsprache und zur Deutschen Gebärdensprache (DGS) erstellt. Exemplarisch wird nachfolgend die Vorgehensweise am Beispiel der deutschen Lautsprache beschrieben. Für die DGS ist es erforderlich, dass die artikulierten Sprachzeichen in einer geeigneten schriftlichen Form abgebildet werden. In diesem Fall wurde eine sequentielle Glossennotation verwendet. Glossen sind Wörter mit lateinischen Großbuchstaben, welche durch zusätzliche Zeichen ergänzt werden. Dabei wird die Bedeutung einer Gebärde durch das Wort repräsentiert. Die visuell-gestische Gebärdensprache weist neben den offensichtlichen Unterschieden in der Modalität auch einige grammatikalische Besonderheiten auf. Weiterführende Informationen zur Deutschen Gebärdensprache findet man z. B. in [1, 3, 5].
Für die Erstellung von Anwendungen für die semantische Sprachverarbeitung stehen im Idealfall geeignete große (semantisch annotierte) Sprachdatenbasen zu Verfügung. Damit können z. B. probabilistische Übersetzungsmodelle und statistische Sprachmodelle erstellt werden. Zur Deutschen Gebärdensprache gibt es kaum geeignetes Sprachmaterial. Oft wurden die erstellten Gebärdensprach-Korpora auch nur partiell transkribiert.
Für die Modellierung semantischer Grammatiken ist in diesem Fall eine eigene Sammlung von Sprachäußerungen, z. B. mit Hilfe von *Wizard-of-Oz-Experimenten*, sinnvoll. Anhand solcher Sprachdaten wurde bereits in [2, 6] die Konstruktion von Transduktoren zur semantischen Interpretation untersucht. In der Domäne D des sprachverarbeitenden Systems ist eine endliche Men-

ge von sprachlichen Realisierungen $\mathcal{U}(D)$ denkbar. Diese werden durch die *lokale Grammatik* beschrieben. Eine *mikrolokale Grammatik* ist ein Teil dieser *lokalen Grammatik* und gehört zu einem *semantischen Schema*. Eine Grammatik, welche nur mögliche Folgen von sprachlichen Zeichen zu einer konkreten Bedeutungsstruktur akzeptiert, wird hier als *Elementargrammatik* bezeichnet. Weiterführende Informationen zu dieser Grammatikhierarchie finden man u. a. in [8]. Nachfolgend wird prinzipiell erörtert, wie aus einer *Elementargrammatik* ein *Äußerungs-Bedeutungs-Transduktor* erzeugt werden kann. Dieses Modellierungsergebnis soll dabei hier die folgenden 2 wesentlichen Funktionen erfüllen:

1. Akzeptor für mögliche Sprachäußerungen der *mikrolokalen Grammatik*
2. Generator einer Struktur von semantischen Konzepten.

Für die Modellierung wurde eine konkrete Situation angenommen, in der verschiedene sprachliche Realisierungen zum Erreichen eines Kommunikationsziels möglich sind. Als Vertreter der zu erstellenden *Elementargrammatik* dient der Satz: *"Dort drüben steht Peter und wartet."*. Die semantischen Einheiten, die sich aus dieser Sprachäußerung ableiten lassen, sind die *relative Ortsangabe (dort drüben)*, die *Handlungen (stehen* und *warten)* sowie der *Vorname einer Person (Peter)*.
Zu dieser Sprachäußerung sind entsprechende *synonymische Syntagmen* zu finden. *Synonymische Syntagmen* sind bedeutungsgleiche Sprachäußerungen, die sich auf dieselbe außersprachliche Bedeutungsstruktur beziehen. Durch eigene Überlegungen sind 96 lautsprachliche Benutzeräußerungen für die *Elementargrammatik* erstellt worden. Parallel dazu wurden passende Gebärdensprachäußerungen von der Fachgruppe Gebärdensprachdolmetschen an der Westsächsischen Hochschule Zwickau artikuliert und transkribiert. Für die Modellierung der gebärdensprachlichen *Elementargrammatik* sind 11 Glossentranskriptionen verwendet worden. Eine direkte Übersetzung der erkannten Sprachzeichenfolge (beschriftete Totalordnung) in eine *Merkmal-Werte-Relation* (beschriftete partielle Ordnung) ist mit endlichen Transduktoren nicht möglich. Die Bedeutungsstruktur wird deshalb zunächst wie in [9] durch eine *MWR-Zeichenkette* repräsentiert. Die dabei verwendete Klammerstruktur definiert den Aufbau der jeweiligen Konzepte und die Relationen der semantischen Einheiten untereinander. Aus diesen Symbolfolgen können in einem Nachbearbeitungsschritt die *Merkmal-Werte-Relationen* erzeugt werden. Die Erstellung eines Transduktors zur *Elementargrammatik* ist durch folgende Schritte realisiert worden:

1. Erstelle zu jeder Sprachäußerung einen endlichen Transduktor T_{Si}, welcher zu der akzeptierten Zeichenfolge eine *MWR-Zeichenkette* generiert.
2. Erzeuge aus der Vereinigung dieser N Transduktoren T_{Si} einen Gesamtautomaten $T_{EG} = \bigoplus_{i=1}^{N} T_{Si}$.
3. Generierung eines *äquivalenten*, jedoch topologisch optimierten Transduktors (optional).

Anschließend kann aus dem Transduktor T_{EG} zur *Elementargrammatik* durch das Einsetzen von *Platzhaltern* ein *Äußerungs-Bedeutungs-Transduktor* zur *mikrolokalen Grammatik* erzeugt werden. Die variablen Parameter der *mikrolokalen Grammatik* können z. B.

- unterschiedliche Datenwerte eines Merkmals,
- verschiedene Ausprägungen von Attributen mit dazugehörigen Werten oder
- ganze untergeordnete *mikrolokale Grammatiken*

sein. Der Grad der Abstraktion wird individuell bei der Modellierung des sprachverstehenden Systems festgelegt. Bei der Erstellung des *Äußerungs-Bedeutungs-Transduktors* zum vorgestellten *semantischen Schema* wurden sowohl für die deutsche Lautsprache als auch für die

Deutsche Gebärdensprache alle 3 Attribute durch *Platzhalter* ersetzt. So sind Generalisierungen für die ausführbaren Aktionen, relative Ortsangaben sowie für verschiedene Namensbezeichnungen von Personen modelliert worden. Die Platzhalter wurden so gestaltet, dass diese vielfältig verwendbar sind. Beispielsweise wurden im Automaten für die verschiedenen Handlungen ein Großteil möglicher Verbflexionen berücksichtigt. Dadurch erhöht sich gleichzeitig die Robustheit der semantischen Interpretation in sprachverarbeitenden Systemen.
Exemplarisch soll nachfolgend kurz der Platzhalter *"name"* für die Person beschrieben werden. Der amtliche Name einer Person kann aus Vornamen und Nachnamen zusammengesetzt sein. Hierbei sollen jeweils auch mögliche Doppelnamen Berücksichtigung finden. Für die Beispielanwendungen wurden je 3 Vor- und Nachnamen verwendet, wobei nur die Sprachzeichen *„Sabine"* und *„Peters"* einem Merkmal eindeutig zugeordnet sind. Im Gegensatz dazu können die genutzten Namen *„Steffen"* und *„Peter"* Vor- oder Familienname sein. Auch Platzhalter be-

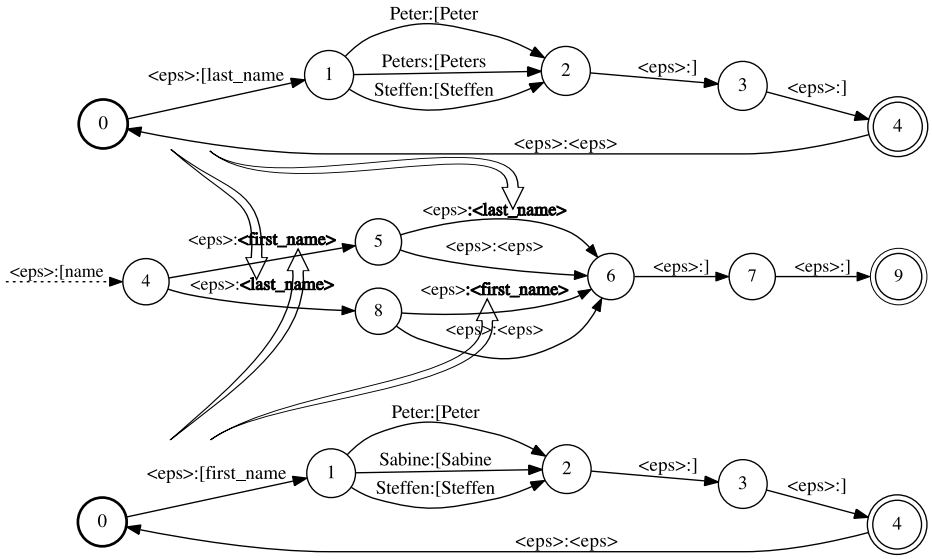

Abbildung 2 - Transduktor für Familiennamen (oben); Ausschnitt des initialen Namenstransduktors mit Platzhaltersymbolen <last_name> und <first_name> (Mitte); Transduktor für Familiennamen (unten)

sitzen u. U. hierarchische Strukturen. In dem initialen Transduktor für den Platzhalter *"name"* sind hier jeweils die untergeordneten Platzhaltersymbole für Vornamen und Familiennamen sowie zusätzliche optionale Hinweiswörter zum Attribut enthalten. So werden in dieser Generalisierung die Sprachzeichen *„Herr"* und *„Frau"* sowie die bestimmten Artikel *„der"* und *„die"* berücksichtigt. Die in der Abbildung 2 visualisierten Subtransduktoren für die deutsche Lautsprache werden dann in den initialen Transduktor *"name"* eingefügt.
In den Transduktor T_{EG} Nachdem alle Platzhalter modelliert worden sind, können diese in den Transduktor T_{EG} eingesetzt werden. Dazu wurden in diesem Automaten zuvor die Ausgabesymbole der semantischen Konzepte durch das dazugehörige Symbol des Platzhalters substituiert. Gleichzeitig sind die zu den Konzepten assoziierten Eingabewörter der Quellsprache durch das leere Symbol ersetzt worden. Danach liegt für beide Sprachformen jeweils ein *Äußerungs-Bedeutungs-Transduktor* T_{UM} vor, welcher für die semantische *Interpretation* und die *Artikulation* verwendet wird.

3 Semantische Interpretation

Ausgangspunkt ist ein Transduktor T_W mit der erkannten Wortfolge einer Sprachäußerung. Dies kann selbstverständlich auch ein gewichteter Worthypothesegraph aus einem realen Spracherkenner sein. Als Beispiel dient nachfolgend der lautsprachliche Satz *"Wartend steht Peter dort drüben."*. In Abbildung 3 ist der dazugehörige Automatengraph dargestellt.

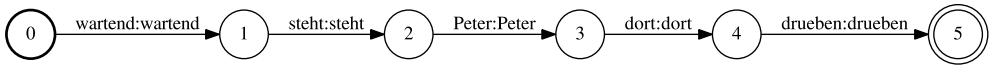

Abbildung 3 - Transduktor T_W für die ausgewählte Wortfolge

Die semantische Interpretation wird durch die Komposition

$$T_{MWR} = T_W \circ T_{UM} \tag{1}$$

mit dem *Äußerungs-Bedeutungs-Transduktor* T_{UM} erreicht. Als Ergebnis erhält man den Transduktor T_{MWR}, welcher für den interpretierten Beispielsatz zwei durchgehende Pfade enthält (vgl. Grafik 4). Die dazugehörigen *MWR-Zeichenketten* sind:

*FVR[PETER[action[wait]][action[stand]]**[person[name[first_name[Peter]]]]***
[loc[rel[there]]]].
*FVR[PETER[action[wait]][action[stand]]**[person[name[last_name[Peter]]]]***
[loc[rel[there]]]].

Diese Sequenzen unterscheiden sich lediglich durch die beiden Interpretationsmöglichkeiten des Sprachzeichens *„Peter"*. In der Beispielanwendung wird zufällig eine Ausgabezeichenfol-

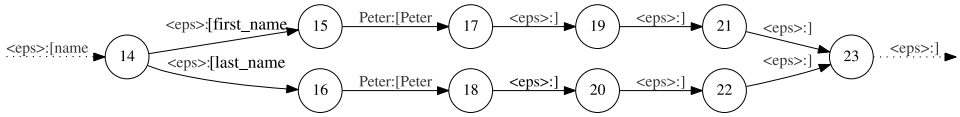

Abbildung 4 - Interpretationsergebnis - Ausschnitt des Transduktorgraphen T_{MWR}

ge ausgewählt und in eine *Merkmal-Werte-Relation* überführt. Die Grafik 5 zeigt die als Baumstruktur visualisierte ungewichtete MWR. Bei der Erzeugung der *Merkmal-Werte-Relation* erfolgt nach [3, 9] eine Kaskadierung gleichnamiger Geschwisterknoten. Damit kann die Reihenfolgeinformation auch in einer Halbordnung berücksichtigt werden.

4 Artikulation mit inversen Äußerungs-Bedeutungs-Transduktoren

Für die Sprachgenerierung werden ebenfalls die modellierten *Äußerungs-Bedeutungs-Transduktoren* verwendet. Eine Generierung von Sprachäußerungen aus einer gegebenen Bedeutungsstruktur erfolgt durch die folgenden 3 Schritte:

1. Erzeugung des Transduktors T_{MWP}, der alle erlaubten Konzeptsequenzen zu einer gegebenen *Merkmal-Werte-Relation* generiert.
2. Zuordnung möglicher Wortfolgen der Zielsprache zu den akzeptierten *MWR-Zeichenfolgen*
3. Zufällige Selektion einer generierten Sprachäußerung. Optional erfolgt die Auswahl aufgrund einer Bewertung des erzeugten Worthypothesegraphen durch ein Sprachmodell.

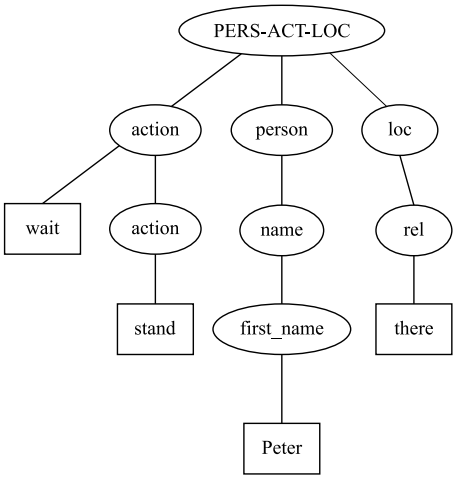

Abbildung 5 - Visualisierte Merkmal-Werte-Relation als Ergebnis des Interpretationsprozesses zur Sprachäußerung *"Wartend steht Peter dort drüben."*

Im ersten Schritt wird eine Linearisierung der Bedeutungsrepräsentation vorgenommen. Dazu wird beginnend mit den Datenwerten (Blätter des Baumes) der gerichtete Graph rekursiv durchsucht. Gleichzeitig werden die Beschriftungen der untersuchten Knoten eingesammelt. In jeder Hierarchiestufe erfolgt eine Permutation der voneinander unabhängigen Geschwisterknoten. Eine Besonderheit stellen die bei der Erzeugung der MWR kaskadierten gleichnamigen Kindknoten dar. Diese sind solange an die nächsthöhere Hierarchieebene zu übergeben, bis keine gleichnamigen Elternknoten mehr auftreten. Erst auf dieser Ebene erfolgt dann die Permutation mit den ermittelten Beschriftungsfolgen der anderen Geschwisterknoten. Dabei werden Sequenzen nicht in die Ergebnisliste aufgenommen, bei der die ursprüngliche Reihenfolge der kaskadierten Elemente nicht gegeben ist. Für die betrachtete Beispieläußerung sind dies also diejenigen beschrifteten *Totalordnungen*, bei der das Konzept $[action[stand]]$ vor $[action[wait]]$ auftritt. Auf diese Weise erhält man zu der *Merkmal-Werte-Relation* des interpretierten Beispielsatzes 12 mögliche *MWR-Zeichenketten*. Diese sind so zu segmentieren, dass die separierten Zeichen Elemente des Ausgabealphabets der *Äußerungs-Bedeutungs-Transduktoren* T_{UM} sind. Anschließend kann aus den segmentierten *MWR-Zeichenketten* der Transduktor T_{MWP} gebildet werden. Als Ergebnis des Übersetzungsprozesses

$$T_{AW} = T_{MWP} \circ T_{UM}^{-1} \{\circ T_{SM}\} \tag{2}$$

erhält man den Automaten T_{AW} mit möglichen Sprachzeichenfolgen der Zielsprache. Der für die Komposition verwendete *inverse Äußerungs-Bedeutungs-Transduktor* T_{UM}^{-1} generiert die zu den akzeptierten Konzeptsequenzen dazugehörigen Wortfolgen.
Zu einem Merkmal-Wert-Paar gehören u. U. mehrere mögliche Wörter bzw. Phrasen. Aufgrund der universalen Modellierung der Platzhalter wird durch die kombinatorische Explosion bei dem Sprachgenerierungsprozess eine große Anzahl unterschiedlicher Wortfolgen erzeugt. Zu der in Bild 5 dargestellten Bedeutungsrepräsentation werden durch den *inversen Äußerungs-Bedeutungs-Transduktor* insgesamt 57024 Sprachzeichensequenzen erzeugt. Eine Vielzahl dieser kombinatorisch möglichen Wortsequenzen sind jedoch grammatikalisch unstimmig.
Aus diesem Grund erfolgte für die deutsche Lautsprache eine gewichtete Bewertung der generierten Wortfolgen. Zu diesem Zweck wurden in den eigenen Untersuchungen

zwei unterschiedliche N-Gramm-Sprachmodelle verwendet. Zum einem wurde anhand einer selbst erstellten Lernstichprobe der Transduktor T_{SM} erzeugt, welcher in diesem Fall ein "mikrolokales" Bigramm-Sprachmodell repräsentiert. Die Lernstichprobe umfasst lediglich 440 kurze Sprachäußerungen, wobei vorwiegend die Lexikoneinträge der *mikrolokalen Grammatik* genutzt wurden. Bei der Erstellung des Bigramm-Sprachmodell wurde ein *Add-One-Smoothing* verwendet.
Zusätzlich erfolgte auch eine Bewertung durch das *Microsoft Web N-Gramm Sprachmodell*, welches aus realen, vorwiegend jedoch englischsprachigen Texten von vielen verschiedenen Webseiten erstellt wurde. Dieses Modell kann auch für die deutsche Sprache verwendet werden. Allerdings ist keine explizite Sprachauswahl bei der Dienstanfrage möglich und auch über den tatsächlichen Umfang der deutschsprachigen Texte im Korpus ist nichts bekannt. Der durch das Sprachmodell gewichtete Transduktor T_{AW} enthält alle möglichen Wortfolgen. Die Gesamtkosten W einer generierbaren Zeichenfolge ergibt sich aus dem dazugehörigen Pfadgewicht. Dieses wird durch die $\otimes$-Operation der beteiligten Kantengewichte $w(e_i)$ berechnet. Kürzere Sprachäußerungen werden dadurch allerdings bevorzugt. Um Wortfolgen mit unterschiedlicher Länge L vergleichen können, ist also eine Normierung des Gesamtgewichts notwendig. Zur Auswertung werden die normierten Kosten W_{rate} für das eigene Bigramm-Modell und für das *Microsoft Sprachmodell* durch

$$W_{rate} = \frac{W}{L} \quad (\text{mit } L > 0) \tag{3}$$

berechnet. Symbole zur Markierung von Anfang und Ende einer endlichen Wortfolge werden bei der Bestimmung der Länge L nicht berücksichtigt.
Die Sprachmodellbewertungen zu der in Bild 5 dargestellten Bedeutungsrepräsentation liefert die in Tabelle 1 gezeigte Anzahl an grammatikalisch unstimmigen Sprachäußerungen. Dabei ist

Sprachmodell	$N = 5$	$N = 10$	$N = 20$	$N = 30$
Bigramm (normiert)	0	2	6	12
Bigramm (gesamt)	2	4	9	15
Microsoft Sprachmodell (normiert)	5	10	19	27
Microsoft Sprachmodell (gesamt)	5	8	17	27

Tabelle 1 - Anzahl grammatikalisch unstimmiger Sätze in der N-besten Liste

für diese einzelne Stichprobe das Ergebnis der Bewertung durch das *Microsoft Web N-Gramm Sprachmodell* unzureichend. Möglicherweise ist dieses allgemein ungeeignet für die Bewertung von deutschen Sprachäußerungen. Bei den normierten Kosten treten bei den bestbewerteten Sequenzen auffällig viele Wortfolgen mit dem auch im Englischen verwendeten Sprachzeichen *„stand"* sowie mit den bestimmten Artikeln (*„der"*, *„die"* und *„das"*) auf. Die Ergebnisse des einfachen "mikrolokalen" Bigramm-Sprachmodells, welches jedoch u. a. auch mit akzeptierten Sätzen der *mikrolokalen Grammatik* trainiert wurde, sind für dieses Beispiel besser.

5 Ausblick

Für weitere Untersuchungen ist die Verwendung mehrerer *mikrolokaler Grammatiken* zu einer Domäne notwendig. Weiterhin ist die Erstellung eines geeigneten größeren Sprachmodells für alle Zielsprachen sinnvoll. Problematisch in diesem Zusammenhang ist die Verfügbarkeit und Generierung von Gebärdensprachdaten. Für die Modellierung der Grammatiken und für

eine Annotation von Sprachdaten ist zudem eine Kooperation mit Gebärdensprachexperten erforderlich. Mit der Definition von *Petrinetz-Transduktoren* (PNT) [4] und elementarer PNT-Operationen sind die notwendigen Grundlagen für eine einheitliche Semantikverarbeitung geschaffen. *Petrinetz-Transduktoren* können als *Äußerungs-Bedeutungs-Transduktoren* und auch zur Bedeutungsrepräsentation verwendet werden.

Literatur

[1] EICHMANN, H., M. HANSEN und J. HESSMANN: *Handbuch deutsche Gebärdensprache: Sprachwissenschaftliche und anwendungsbezogene Perspektiven*, Bd. 50 d. Reihe *Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser*. Signum, Seedorf, 2012.

[2] HUBER, M., C. KÖLBL, R. LORENZ und G. WIRSCHING: *Konstruktion von UMP-Transduktoren aus Wizard-of-Oz Daten*. In: WAGNER, P. (Hrsg.): *Elektronische Sprachsignalverarbeitung 2013*, Bd. 65 d. Reihe *Studientexte zur Sprachkommunikation*, S. S. 111 – 118. TUDpress, Dresden, 2013.

[3] LINDEMANN, J.: *Semantische Verarbeitung von Gebärdensprache in intelligenten hierarchischen Sprachdialogsystemen*. In: HOFFMANN, R. (Hrsg.): *Elektronische Sprachsignalverarbeitung 2014*, Bd. 71 d. Reihe *Studientexte zur Sprachkommunikation*, S. 118–125. TUDpress, Dresden, 2014.

[4] LORENZ, R., M. HUBER und G. WIRSCHING: *On Weighted Petri Net Transducers*. In: *Application and Theory of Petri Nets and Concurrency*, S. 233–252. Springer, 2014.

[5] PAPASPYROU, C., A. VON MEYENN, M. MATTHAEI und B. HERRMANN: *Grammatik der deutschen Gebärdensprache aus der Sicht gehörloser Fachleute*. Signum, Seedorf, 2008.

[6] RITTER, D. und G. WIRSCHING: *Konstruktion einer mikrolokalen Grammatik mit OpenFST am Beispiel einer Home-Entertainment-Anwendung*. In: HOFFMANN, R. (Hrsg.): *Elektronische Sprachsignalverarbeitung 2014*, Bd. 71 d. Reihe *Studientexte zur Sprachkommunikation*. TUDpress Verlag der Wissenschaften Dresden, Dresden, 2014.

[7] WIRSCHING, G. und C. KÖLBL: *Language Modeling with Utterance-Meaning-Pairs. Technical Report 2011-12, Institute of Computer Science, University of Augsburg*, 2011.

[8] WIRSCHING, G. and R. LORENZ: *Towards meaning-oriented language modeling*. In *Cognitive Infocommunications (CogInfoCom), 2013 IEEE 4th International Conference*, pp. 369–374. 2013.

[9] WIRSCHING, G. und M. WOLFF: *Semantische Dekodierung von Sprachsignalen am Beispiel einer Mikrofonfeldsteuerung*. In: HOFFMANN, R. (Hrsg.): *Elektronische Sprachsignalverarbeitung 2014*, Bd. 71 d. Reihe *Studientexte zur Sprachkommunikation*, S. 104–109. TUDpress Verlag der Wissenschaften Dresden, Dresden, 2014.