

*“Ja zur Not geht’s auch am Samstag”*  
**oder**  
**Wie die Prosodie der Linguistik die Satzzeichen  
verschaffen kann**

Elmar Nöth

Universität Erlangen-Nürnberg, Lehrstuhl für Mustererkennung (Informatik 5),  
Martensstr. 3, 91058 Erlangen, Germany

ABSTRACT

In gesprochener Sprache werden prosodische Mittel u.a. dazu benutzt, um

- Phrasengrenzen
- Phrasenakzente
- Satzmodus

zu markieren. Dadurch wird die Äußerung strukturiert und für den Hörer die Bedeutungsanalyse erleichtert. Die prosodische Markierung von Phrasengrenzen und Satzmodus ersetzt sozusagen die Satzzeichen in geschriebener Sprache.

Die Arbeiten, über die hier berichtet wird, sind in das VERBMOBIL-Projekt eingebunden. Im Anwendungsgebiet von VERBMOBIL – Simultanübersetzung von Terminabsprache-Dialogen – müssen auch typisch spontansprachliche Phänomene modelliert werden, die bei gelesener und bei geschriebener Sprache nicht auftreten. Insbesondere treten bei Spontansprache häufig mehrere kurze, elliptische satzwertige Ausdrücke innerhalb einer Äußerung auf (So bestanden in einem Korpus von 50 VERBMOBIL-Dialogen 70 % der Äußerungen aus mehr als einem Satz).

Hierdurch ergibt sich für die linguistische Analyse das Problem, daß zwischen vielen kombinatorisch möglichen Parses disambiguiert werden muß. In VERBMOBIL wird prosodische Information letztendlich benötigt für:

- Wahl der korrekten Übersetzungsalternative
- Wahl der korrekten prosodischen Form bei der Synthese.

Ein Beispiel aus dem VERBMOBIL-Anwendungsgebiet soll die Problematik verdeutlichen:

Ja / zur Not / geht’s / auch / am Samstag

Hier kann es sich um einen, zwei, drei oder gar vier satzwertige Ausdrücke handeln. Die folgende Tabelle zeigt die möglichen syntaktischen Konstrukte, die zu mindestens 36 möglichen Parses führen:

		Funktion	Satzzeichen
ja	1	ell. Aussagesatz	.
	2	(ell. (Rück)fragesatz/ diskurssteuernde Gliederungspartikel	?
ja ...	3	Satzaufaktartikel	
zur Not	1	ell. Aussagesatz	.
	2	ell. (Rück)fragesatz	?
zur Not...	3	Satzadverbial	
geht's	1	Fragesatz	?
...geht's	2	Ende eines Aussagesatzes	.
...auch	1	Ende eines Fragesatzes	?
	2	Ende eines Aussagesatzes	.
am Samstag	1	ell. Aussagesatz	.
	2	ell. Fragesatz	?

In diesem Vortrag werden Ergebnisse darüber präsentiert, wie in einem Bottom-up-Ansatz prosodische Information dem Parser zugänglich gemacht werden kann. Ausgehend vom Worthypothesengraphen wird für jedes Wort des Graphen die Wahrscheinlichkeit berechnet, daß nach dem Wort eine prosodisch markierte Grenze vorliegt. Die unterschiedlichen benutzten Merkmale und Klassifikatoren werden vorgestellt. Diese Wahrscheinlichkeiten werden in den Graphen eingetragen.

In die Grammatik wurde eine spezielle syntaktische Kategorie "break" eingeführt, die in den Phrasenstrukturregeln ähnlich wie Satzzeichen verwendet werden kann. Der Syntaxparser benutzt die prosodische Information, um

- die richtige Wortkette zu finden
- und für die richtige Wortkette
- die möglichen Parses zu reduzieren.

Da zur Zeit noch nicht genügend adäquat gelabeltes VERBMOBIL-Material vorhanden ist, werden Ergebnisse für eine Teilmenge des ERBA-Korpus präsentiert. Bei den untersuchten Sätzen handelt es sich um gelesene Hauptsätze mit infinitivem Nebensatz. Es zeigt sich, daß Grenzen zwischen Haupt- und Nebensatz gut erkannt werden. Wird die richtige Wortkette im Graphen gefunden, so führt die Beachtung der prosodischen Grenzhypothesen zu einer Reduktion der möglichen Parses um 70 %.

Im Ausblick wird eingegangen auf zusätzliche Probleme, die sich beim Übergang von gelesenem auf spontanes Sprachmaterial (VERBMOBIL-Stichprobe) ergeben.