

Hagen Hirschmann (Berlin, Deutschland)

Von der Restkategorie Adverb zur korpusrelevanten syntaktischen Ausdifferenzierung

Abstract

Der Vortrag behandelt die funktional-syntaktische Ausdifferenzierung der im Stuttgart-Tübingen Tagset (STTS) vorgeschlagenen Wortartenkategorie Adverb (vertreten durch das Tag "ADV") und legt anhand einer Fallstudie die Relevanz einer solchen Ausdifferenzierung für linguistische Fragestellungen, welche anhand von flach annotierten Korpusdaten behandelt werden, dar.

Die Betrachtungen bezüglich des STTS können auch auf andere Wortarten-Tagsets für deutschsprachige Textkorpora angewandt werden, denn kein allgemein gebräuchliches Tagset liefert bislang eine differenzierte, kriterienreine Klassifikation für den Wortartenbereich Partikel/Adverb.

Die Wortartenkategorie "ADV" im STTS umfasst eine morphologisch homogene (alle Elemente sind unflektierbar), jedoch syntaktisch äußerst heterogene Klasse von Wörtern.

Dies entspricht dem Granularitätsgrad, welches die Wortart Adverb im traditionellen Wortartensystem der 10 Wortarten (Substantiv, Verb, Adjektiv, Artikel, Pronomen, Adverb, Konjunktion, Präposition, Numerale, Interjektion) hat.

Dieser Vortrag begründet die Notwendigkeit einer Subklassifikation in diesem Bereich für deutsche Wortartentagsets. Mit dem Begriff 'Adverb' (in einfachen Anführungszeichen) wird im Folgenden auf den Gesamtbestand von Adverbien und Partikeln Bezug genommen.

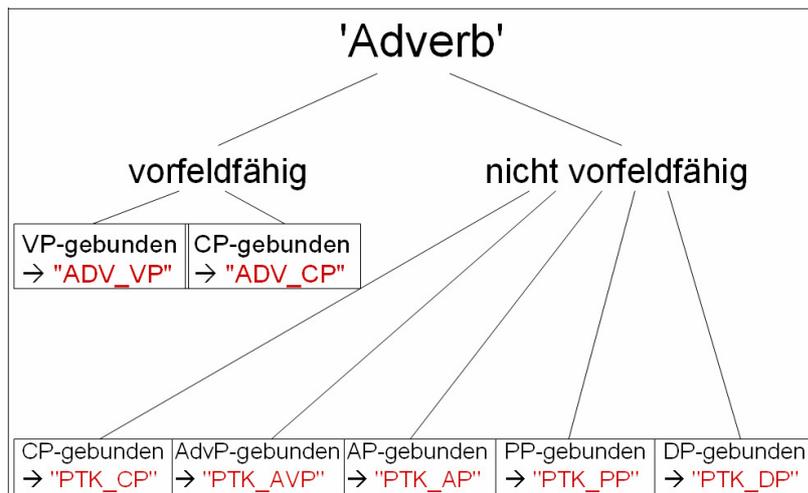
Die Wortart 'Adverb' und deren Unterklassifizierung ist mit Sicherheit die umstrittenste Wortartkategorie (vgl. Hoffmann 2007, S. 227ff). Zunächst werden die Entwicklung und aktuelle Strömungen in der Behandlung dieser Klasse in der germanistischen Linguistik dargestellt. Hierbei wird das auf das Problem der wissenschaftlichen Kategorisierung und der Kriterienreinheit in der Wortartenklassifizierung allgemein und bei den 'Adverbien' im Speziellen eingegangen. Daran anknüpfend wird an Korpusanfragen und Treffervergleichen die Problematik der Wortart "ADV" im STTS dargestellt.

Eine Suche im STTS-getaggten Korpus nach bestimmten 'Adverb'-Klassen ist ohne weiteres nicht möglich. Aus diesem Grund haben bisherige korpusbasierte Studien vor allem zu bestimmten Partikelklassen ausschließlich oberflächenbasiert, also mittels Suche nach bestimmten Lexemen, arbeiten können (vgl. z.B. Thurmair 1991 oder Lemnitzer 2001).

Wie gezeigt wird, erlaubt ein syntaktisch annotiertes Korpus eine Annäherung an bestimmte syntaktisch-funktionale 'Adverb'-Klassen wie Fokus- oder Modalpartikeln. Exemplarisch werden hierfür Suchen im Tiger-Korpus illustriert und gezeigt, wie auf diese Weise Wortlisten für die einzelnen Klassen generiert werden können und mit bestehenden, auf Tradition oder Intuition beruhenden Wortlisten kontrastiert werden können.

Leider gibt es für das Deutsche nur wenig syntaktisch annotierte Korpora, weil automatisches Parsing noch nicht ohne zu hohe Fehlerquoten möglich ist und manuelles Annotieren äußerst zeitaufwändig ist. Selbst die beiden "großen" syntaktisch annotierten Korpora – das Tiger-Korpus und die TüBa-D/Z-Baumbank sind für viele Studien recht klein und bestehen ausschließlich aus Zeitungssprache.

Um dem entgegenzuwirken und auch in nicht syntaktisch annotierten Korpora unterschiedliche Klassen von 'Adverbien' zu finden, wird eine korpusrelevante funktional-syntaktische Ausdifferenzierung der 'Adverbien' nach zwei hierarchisch angeordneten Kriterien und damit verbunden die Subklassifikation des STTS-Wortartentags "ADV" vorgestellt:



Tagset für die Subklassifizierung des Wortarttags "ADV" nach dem STTS. Der Begriff "vorfeldfähig" wird hier synonym zu "funktionsfähig" verwendet.

Eine Fallstudie zu dem Gebrauch unterschiedlicher Adverbien und Partikeln und deren Abfolge bei fortgeschrittenen Lernenden des Deutschen als Fremdsprache zeigt die Relevanz einer solchen syntaktischen Kategorisierung im Bereich von Spracherwerbsstudien.

Die Fallstudie zeigt unter anderem, dass sich durch diese Annotation selbst in flach annotierten Korpora nach komplexen Adverbialen bzw. Adverbialphrasen (*Und gewinnen will sie [immer wieder].*) suchen lässt und sich diese Einheiten von Adverbienketten, die getrennte Phrasen bilden (*Sie hofft, dass das [morgen] [so] klappt.*), trennen lassen und somit Aussagen über die Produktivität und Variabilität von syntaktischen Strukturen möglich sind.

Eine Erweiterung des etablierten STTS-Tagsets wie die vorgeschlagene ist die einzige Möglichkeit, in flach annotierten Korpora Studien zu bestimmten Klassen von 'Adverbien' durchzuführen und kann erheblich zur weiteren Theoriebildung in diesem Bereich beitragen.

Literatur

- Admoni, Wladimir (1980): Der deutsche Sprachbau, München; Beck.
- Dimroth, Christine (2002): Topics, assertions, and additive words: how L2 learners get from information structure to target language syntax. In: *Linguistics* 40, 891-923.
- Ferraresi, Gisella (2006): Modalpartikeln und Satzadverbale – Gemeinsamkeiten und Unterschiede. In: *Deutsch als Fremdsprache* 43, 208-214.
- Helbig, Gerhard; Buscha, Joachim (2001): Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht, Leipzig, Berlin, München; Langenscheidt.
- Hoffmann, Ludger (Hg. 2007): Handbuch der deutschen Wortarten, Berlin; New York; Walter de Gruyter.
- Lemnitzer, Lothar (2001): Wann kommt er denn nun endlich zur Sache? Modalpartikel-Kombinationen – Eine korpusbasierte Untersuchung. In: Lehr, Andrea et al. (2001): *Sprache im Alltag. Beiträge zu neuen Perspektiven in der Linguistik*, 349-371, Berlin; New York; Walter de Gruyter.
- Stuttgart-Tübingen-Tagset
<http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.html>
- Thurmair, Maria (1991): Zum Gebrauch der Modalpartikel *denn* in Fragesätzen. Eine korpusbasierte Untersuchung. In: E. Klein / F.P. Duteil / K.H. Wagner (Hgg.): *Betriebslinguistik und Linguistikbetrieb. Akten des 24. Linguistischen Kolloquiums, Universität Bremen, 4.-6. September 1989*. Tübingen; Niemeyer, 377-388.
- Zifonun, Gisela; Hoffmann, Ludger; Strecker, Bruno (1997): *Grammatik der deutschen Sprache*. Band 1. Berlin/New York; Walter de Gruyter.