gefördert vom
Bundesministerium
für Forschung
und Technologie

Verbmobil
Verbundvorhaben

Spontansprachliche syntaktische Phänomene: Analyse eines Korpus aus der Domäne „Terminabsprache“

Herbert S. Tropf

Siemens AG, München

Vm
Bericht 12
April 1994

April 1994

Herbert S. Tropf

Siemens AG
ZFE ST SN 54
D–81730 München

Tel.: (089) 636 - 44195
e-mail: `Herbert.Tropf@zfe.siemens.de`

Gehört zum Antragsabschnitt: 6.5 Spontansprachliche Konstruktionen

Das diesem Bericht zugrundeliegende Forschungsvorhaben wurde mit Mitteln des Bundesministers für Forschung und Technologie unter dem Förderkennzeichen 01 IV 102 A/0 gefördert. Die Verantwortung für den Inhalt dieser Arbeit liegt bei dem Autor.

Inhaltsverzeichnis

Tabellenverzeichnis

1 Einleitung

Der angestrebte Einsatz eines Verbmobil–Gerätes impliziert eine spezifische Sprechsituation: Es ist eine Mensch–zu–Mensch–Kommunikation, die durch ein bei Bedarf manuell zu aktivierendes Übersetzungsgerät unterstützt wird. Insofern unterscheidet sich die Kommunikationsart sowohl von der „normalen“ Mensch–zu–Mensch–Kommunikation, wie sie bei Alltagsgesprächen üblich ist, als auch von der Mensch–Maschine–Kommunikation, wie sie etwa in der Verwendung von Datenbankabfragesystemen oder Beratungssystemen vorkommt.[1]

Drei Faktoren sind für das Sprachverhalten des Benutzers beim anvisierten Verbmobil–Einsatz kennzeichnend:

- Es handelt sich um gesprochene und nicht um geschriebene Sprache.
- Es handelt sich um spontan gesprochene Sprache, die grammatisch und lexikalisch unrestringiert ist — dies im Gegensatz zu vorgelesenem oder auswendig gesprochenem Material, wie es bei existierenden Demonstrationssystemen für kontinuierlich gesprochenen Input gegenwärtig noch üblich ist.
- Das Übersetzungsgerät muß vom Sprecher bewußt kontrolliert, d.h. durch Tastenbetätigung aktiviert bzw. deaktiviert werden, sobald die Übersetzung einer Passage begonnen bzw. beendet werden soll.

Diese Faktoren lassen erwarten, daß bei einer tatsächlichen Verbmobil–Verwendung sowohl sprachliche Phänomene auftreten, wie sie in der Literatur über normale spontansprachliche Mensch–zu–Mensch–Kommunikation

[1] Für ein Datenbankabfragesystem mit kontinuierlich gesprochener Sprache vgl. z.B. SPICOS [Niedermair et al. 1990]) und für ein Beratungssystemen mit geschriebenem Input vgl. z.B. WISBER [Horacek et al. 1988]).

beschrieben werden, als auch Phänomene, die für *computer talk* charakteristisch sind.[2]

Das Datenmaterial, das Gegenstand der vorliegenden Analyse ist, ist das Ergebnis von mehreren Dialog–Sitzungen, bei denen die Sprechsituation des echten Verbmobil–Einsatzes simuliert wurde. Diese Sprechsituation ist gekennzeichnet durch einen unrestringierten Sprachgebrauch und die bewußte Tastenbedienung je Redebeitrag. Eine Übersetzung durch das Gerät wurde bei der Datenaufnahme jedoch nicht simuliert. Bezüglich dieser Differenz werden beim realen Verbmobil–Einsatz zusätzliche spezifische Phänomene erwartet, die jedoch stärker die Dialogebene (z.B. Klärungsdialoge) als die syntaktische Ebene betreffen dürften.

Der vorliegende Bericht gibt neben einer Zielformulierung und Korpusskizze eine quantitative Übersicht über syntaktische Strukturen, die in der für Verbmobil angestrebten Anwendungsdomäne tatsächlich vorkommen. Abschließend werden kurz einige Folgerungen erläutert.

[2] Letzteres bezeichnet eine Art hyperkorrektes Sprachverhalten des Systembenutzers, wie es in einer Mensch–Maschine–Kommunikation typischerweise auftritt (vgl. [Krause und Hitzenberger 1992]).

2 Ziel der Korpusanalyse

Der übliche Gegenstandsbereich syntaktischer Beschreibungen ist traditionell die geschriebene Sprache, und es werden daraus Phänomene ausgewählt, die als Prüfstein der linguistischen Theoriebildung erachtet werden und/oder für eine in der Regel rudimentäre Syntax eines lauffähigen Systems unerläßlich sind. Eine solche Selektion ist jedoch nicht repräsentativ für das, was in den verschiedenen Domänen gesprochener Sprache tatsächlich realisiert wird.

Primäres Ziel der vorliegenden Korpusanalyse ist es, eine quantitative Übersicht über syntaktische Phänomene einer spezifischen Domäne gesproche ner Sprache zu erstellen, die für den Verbmobil–Einsatz unmittelbar einschlägig ist, und zwar hauptsächlich über jene Phänomene, die vom üblichen linguistischen Gegenstandsbereich abweichen. Darüberhinaus gibt eine solche Untersuchung Hinweise darauf, was bei der syntaktischen Verarbeitung gesprochener Sprache generell, also unabhängig von einer spezifischen Domäne, besonders zu berücksichtigen ist.

Diese Übersicht ist nicht nur als Liste der abzudeckenden syntaktischen Strukturen gedacht. Vielmehr können aufgrund der quantitativen Information die Arbeiten priorisiert werden, die für eine umfassende syntaktische Beschreibung in Form eines Regelapparates notwendig sind. Mit einer solchen Priorisierung kann gewährleistet werden, daß zu einem möglichst frühen Zeitpunkt ein relativ hoher Abdeckungsgrad der zu verarbeitenden Strukturen erreicht wird. Dabei kann eine solche Priorisierung durchaus bedeuten, daß ein bestehender Arbeitsplan zu modifizieren ist.

3 Skizze des Korpus

Gegenstand der Analyse ist das transliterierte Korpus der Universität Karlsruhe, Institut LKD, wie es dem Verbmobil–Konsortium im August bzw. November 1993 als kostenloser Service zur Verfügung gestellt wurde.

Das Korpus enthält Spontansprache von jeweils zwei Kommunikationspartnern, die in einer möglichst ungezwungenen Laboratmosphäre anhand einer konkreten Aufgabenstellung elizitiert wurde. Es handelt sich dabei nicht um vollkommen freie Spontansprache, wie sie in Alltagssituationen üblich ist. Die Abweichungen davon betreffen technische Randbedingungen für die Aufnahmen. Inhalt, syntaktische Struktur oder Wortwahl sind jedoch in keiner Weise restringiert.

Domäne

Ziel der Konversation zwischen den Kommunikationspartnern ist die Vereinbarung eines Termins für ein gemeinsames Treffen von 2 Stunden. Jeder der zwei Kommunikationspartner hatte einen unterschiedlich ausgefüllten Terminkalender als Vorlage. Zusätzlich wurde vorgegeben, daß die Dialoge nicht zu kurz (mindestens 6–8 Sätze) sein sollen, und daß je Dialog nur ein Termin zu vereinbaren und danach der Dialog zu beenden ist. Außerdem wurde darauf hingewiesen, daß das Gespräch nicht zu stark thematisch abweichen soll.

Umfang

Das Korpus besteht bisher aus insgesamt 50 Dialogen. Dies sind 641 Redebeiträge[3], die typischerweise je 1 bis 3 Sätze umfassen.

[3] „Redebeitrag“ und „Äußerung“ werden hier synonym verwendet.

Versuchspersonen für alle 50 Dialoge waren 24 Studenten (3 weiblich, 21 männlich).

Datenaufnahme

Die Abweichungen gegenüber vollkommen freier Spontansprache bestehen im wesentlichen darin, daß über Mikrophon (*head–set*) und mit Tastenbetätigung an einem Computer gesprochen werden mußte. Es durfte also jeweils immer nur eine Person sprechen. Die Zeit für einen Redebeitrag war auf 30 Sekunden beschränkt.

Um nonverbale Kommunikation zu vermeiden, wurden die Sprecher aufgefordert, sich nicht anzuschauen. Außerdem wurden sie aufgefordert, Hochdeutsch zu sprechen. Die Versuchspersonen wurden darauf hingewiesen, daß das Datenmaterial für die Entwicklung eines Spracherkennungs– und Übersetzungssystems verwendet wird. Alle Vorgaben wurden den Sprechern vor der Aufnahme schriftlich zur Kenntnis gegeben. Vor den eigentlichen Aufnahmen wurden Probeaufnahmen für Systemtests und Geräteeinstellung gemacht.

Transliteration

Die Transliteration der aufgenommenen Sprachdaten basiert auf der deutschen Standardorthographie. Sie will insbesondere keine phonetische Transkription sein. Über die intra– und intersubjektive Konsistenz der Transliteration liegen keine Angaben vor. Neben dem Wortlaut wurde u.a. Folgendes in die Transliteration aufgenommen:

- Pausen (weniger/mehr als 2 Sekunden)
- mit der menschlichen Stimme erzeugte Geräusche (z.B. Häsitationslaute, Atemgeräusche, Lippenschmatzen, Zungenschnalzen, Lachen)

- nicht mit der menschlichen Stimme erzeugte spezifische Geräusche, deren Unterscheidung für Training und Test von Spracherkennern relevant ist (z.B. Rascheln, Klingeln, Knallen etc.)
- andere Geräusche, deren Unterscheidung für Training und Test von Spracherkennern irrelevant ist
- Markierung abgeschnittener, d.h. nicht aufgenommener Rede
- deutlich gedehnte Aussprache

Annotationen, die über die Transliteration im engeren Sinne hinausgehen, wurden für folgende Phänomene gemacht:

- falsche Anfänge (mit Kennzeichnung des entsprechenden Bereiches)
- falsche Aussprache (möglichst mit Korrektur)
- Ende eines Satzes, eines Gedanken oder eines Fragmentes in einer langen Äußerung
- Intonationsverläufe (terminal, interrogativ, progredient)

Diese Annotationen haben prätheoretischen Status. Sie reflektieren die Intuition des Transliteraten und können somit erste wertvolle Hinweise für eine systematische Interpretation geben.

4 Methode

Aufbauend auf der Literatur zur gesprochenen Sprache bei Mensch–zu–Mensch–Kommunikation (vgl. z.B. [Altmann 1981], [Bayer 1973], [Betten 1976], [Engel 1974], [Höhne–Leska 1975], [Krause und Hitzenberger 1992], [Rath 1975], [Schachtl und Block 1991], [Schank und Schoenthal 1976], [Steger 1987], [Auer 1991]) wurde unter besonderer Berücksichtigung von [Altmann 1981] und der ASL–Studie [Schachtl und Block 1991] ein vorläufiges Klassifikationsschema erstellt. Nicht als Klasse aufgenommen wurden Strukturen, die in satzorientierten Sprachverarbeitungssystemen für geschriebenen Input ohnehin üblich sind.

Daran anschließend wurde dieses Schema aufgrund wiederholter Dateninspektion modifiziert, d.h. Klassen hinzugefügt, gestrichen oder zusammengefaßt. Das modifizierte Schema umfaßt 68 Klassen, die gruppiert und hierarchisch geordnet sind.

Danach wurde die Vorkommenshäufigkeit der einzelnen Klassen im Korpus ermittelt.

5 Ergebnisse

In diesem Abschnitt werden die Ergebnisse der Klassenauszählung dargestellt. Zuerst werden Phänomene der Äußerungsebene erläutert, danach Phänomene der Satzebene.

Bezugsgröße für die Angaben der relativen Häufigkeiten und Mittelwerte sind die 641 Redebeiträge aller 50 Dialoge des Korpus. Von den 68 vorgegebenen Klassen kommen in diesem Korpus 55 tatsächlich vor. Insgesamt handelt es sich um 3141 Belegstellen, die als solche markiert sind.

Wesentlich für das Verständnis des Folgenden ist das Konzept der freien Phrase, wie sie hier festgelegt wird. Diese Klasse umfaßt satzeinleitende Wörter wie in (1), Herausstellungen im Sinne von [Altmann 1981] (ohne Parenthesen), d.h. Linksversetzung, Rechtsversetzung, Ausklammerung, Nachtrag, als auch Strukturen, die als Satzellipsen aufgefaßt werden können. Zu letzteren zählen insbesondere Adjazenzellipsen (vgl. [Klein 1985]) wie in (2–5) als auch Interjektionen, Antwortpartikeln und Gruß–, Dank– und Entschuldigungsfloskeln wie in (6–9).[4]

(1) *<u>also</u> der Montag sieht bei mir schlecht aus.*

(2) *<u>Moment</u>, heute ist der zweite März.*

(3) *<u>nach sechzehn Uhr</u>, besser wär's sechzehn Uhr.*

(4) *es geht nicht zwischen sieben und neun, <u>keine Chance</u>.*

(5) *dann müßen wir ein' anderen Tag vereinbaren, <u>am fünften</u> <u>vielleicht</u>.*

(6) *<u>oh</u>, das is' schlecht.*

[4] Falls nicht besonders vermerkt, sind alle angeführten Beispiele dem Korpus entnommen, allerdings auf das Essentielle reduziert. Die Struktur, um die es geht, ist mitunter zur einfacheren Identifizierung unterstrichen.

(7) *in Ordnung, dann machen wir das so.*
(8) *sorry, aber auch Mittwoch bin ich nicht im Hause.*
(9) *hallo, gut daß ich Sie noch treffe.*

5.1 Äußerungsebene

In Tabelle 1 wird zunächst eine Übersicht über bestimmte Phänomene der Äußerungsebene gegeben. Sie zeigt, daß Äußerungen, die aus genau 1 vollständigen Satz bestehen, also keine Mehrsatz–Äußerungen oder Satzfragmente sind, und außerdem keine freien Phrasen haben, relativ wenig vorkommen (13,0 %). Dies sind Sätze wie (10).

(10) *können wir uns dann nich' auf halb drei einigen?*

Tabelle1: **Phänomene auf Äußerungsebene**

Klasse	% [a]	Ø [b]
Äußerung mit genau 1 vollständigen Satz ohne freie Phrasen	13.0	.130
Mehrsatz–Äußerung oder Satz mit freien Phrasen	70.5	.705
Äußerung ohne vollständigen Satz	16.5	.165
freie Phrase	70.7	1.462
Datums– oder Zeitangabe	63.3	.958
Häsitationspartikel	17.4	.201
Pause	30.7	.420

[a] % = Prozentsatz der Äußerungen mit mindestens 1 Beleg
[b] Ø = Mittelwert für die Anzahl der Belege je Äußerung (bezogen auf alle Äußerungen)

Diese Art von Sätzen, die zugegebenermaßen intern mitunter eine recht komplexe Struktur aufweisen können, ist aber der übliche und meistens der einzig mögliche Input, der von den Grammatiken und Parsern gegenwärtiger sprachverstehender Systeme akzeptiert wird. Insbesondere fehlt die Fähigkeit, Mehrsatz–Äußerungen und Satzfragmente behandeln zu können. Die bisherige Leistungsklasse reicht für ein Übersetzungssystem, das in der für Verbmobil anvisierten Domäne einsatzfähig sein soll, offensichtlich nicht aus.

Dies wird durch komplementäre Angaben aus Tabelle 1 weiter verdeutlicht. Neben den Äußerungen mit genau 1 vollständigen Satz ohne freie Phrasen bestehen 70,5 % der Äußerungen aus einer Mehrsatz-Äußerungen (mit/ohne freie Phrasen) oder aus genau einem vollständigen Satz mit freien Phrasen (vgl. z.B. (11–12)).

(11) *das ist schlecht. da habe ich um vierzehn Uhr einen Termin.*

(12) *ja, nach dem Mittagessen, das ist prima.*

Dies sind alles Äußerungen, die von den Grammatiken und Parsern gegenwärtiger sprachverstehender Systeme üblicherweise nicht verarbeitet werden können. Dies gilt insbesondere für Äußerungen, die nur aus einer oder mehreren freien Phrasen ohne Matrixsatz bestehen (vgl. z.B. (13)).

(13) *ach, Sie schon wieder.*

Insgesamt zeigt sich die quantitative Dominanz der freien Phrasen in dieser Art von Sprachdaten darin, daß in mehr als 70 % aller Äußerungen mindestens eine freie Phrase belegt ist, und daß durchschnittlich auf jede Äußerung ca. 1,5 freie Phrasen kommen.

Nicht überraschend für die Domäne Terminabsprache ist, daß in knapp zwei Drittel aller Äußerungen mindestens 1 Datums– oder Zeitangabe belegt ist (vgl. Tabelle 1). Durchschnittlich kommt auf jede Äußerung knapp 0,958 solcher Angaben. Dies können Standardangaben sein wie in (14–19), informelle Angaben wie in (20–25) und Kombinationen von Datums– und Zeitangabe wie in (26–27)

(14) *Freitag den siebzehnten September*
(15) *sechzehnter März neunzehn hundert drei un' neunzig*
(16) *die zehnte Kalenderwoche*
(17) *in den ersten zwei August–Wochen*
(18) *um sechzehn Uhr vierzig*
(19) *nachmittags zwischen vierzehn und sechzehn Uhr*

(20) *der Donnerstag nach Ostern*
(21) *am kommenden Freitag*
(22) *Ende Mai*
(23) *bis kurz nach zwölf*
(24) *pünktlich acht Uhr*
(25) *so um neun*

(26) *Dienstag früh um zehn*
(27) *Freitag zwölfter März vormittags*

Eine wesentliche Eigenschaft gesprochener Sprache ist das Vorkommen von Häsitationslauten und Pausen. Beide Phänomene müssen von realistischen sprachverarbeitenden Systemen behandelt werden können. Dabei können sie gegebenenfalls ausgenutzt werden zur Desambiguierung, z.B. in der Bestimmung von Satzgrenzen (vgl. (28–30)) oder von Neuansätzen (vgl. (31–33)).

(28) *am besten wär's morgens früh* *silence* *wär' das möglich.*

(29) *dann nachmittags gegen vierzehn Uhr* *pause* *können wir uns da treffen?*

(30) *das Problem ist /eh/ nächste Woche zwischen Dienstag und Donnerstag bin ich nicht hier.*

(31) *Dienstag zur Not ja, obwohl ich da* *pause* *ja Dienstag ist möglich.*

(32) *wie wär's denn am* *pause* *[Jun] am zehnten Juni?*

(33) *ja ich /eh/ von mir aus können wir das machen.*

Die Vorkommenshäufigkeiten zeigen, daß insbesondere die Pause eine potentiell relevante Größe darstellt (vgl. Tabelle 1). In ca. 31 % aller Äußerungen wurde von den Transliteraten jeweils mindestens 1 Pause (meistens kürzer als 2 Sekunden) wahrgenommen, wobei die durchschnittliche Vorkommenshäufigkeit je Äußerung 0,42 beträgt.

Freie Phrase

Da die freie Phrase in der Spontansprache der hier untersuchten Domäne — wie wahrscheinlich generell in der Spontansprache — eine so herausragende Rolle spielt, soll an dieser Stelle etwas detaillierter darauf eingegangen werden. Zunächst wird kurz auf die Positionen der freien Phrase mit den entsprechenden Vorkommenshäufigkeiten eingegangen, dann auf ihre strukturelle Ausprägungen mit den entsprechenden Vorkommenshäufigkeiten.

In ungefähr der Hälfte der Äußerungen steht mindestens 1 freie Phrase links des entsprechenden Matrixsatzes wie z.B. in (34) (vgl. Tabelle 2). Dabei haben durchschnittlich alle Äußerungen jeweils 0,846 freie Phrasen links eines Satzes. Dies ist in den meisten Fällen auch der Äußerungsanfang. In geringerem Maße kommen freie Phrasen rechts eines entsprechenden Matrixsatzes vor (vgl. (35)). In ca. 14 % aller Äußerungen kommt mindestens 1 freie Phrase vor, ohne daß die entsprechende Äußerung überhaupt einen Satz enthält (vgl. (36)).

(34) *höchstens eine halbe Stunde, ab zehn Uhr hab' ich einen Folge-Termin.*

(35) *da findet ein Seminar statt, auf dem ich bin, von neun bis sechzehn Uhr.*

(36) *okay, Dienstag, bei dir.*

Tabelle2: **Position der freien Phrase**

Klasse	%	Ø
Freie Phrase links vom Satz	51.01	.846
Freie Phrase rechts vom Satz	19.03	.304
Freie Phrase ohne Matrixsatz	13.88	.332

Bei Mehrsatz–Äußerungen kommen freie Phrasen allerdings auch äußerungsintern vor. Hier ist mitunter nicht zweifelsfrei zu entscheiden, ob eine solche Phrase dem vorausgehenden oder dem nachfolgenden Matrixsatz zuzuordnen ist. Aufgrund der realisierten Intonation könnten die meisten dieser Fälle klar entschieden werden. Bei alleiniger Vorlage des Transliterationstextes ohne detaillierte intonatorische Information konnte in solchen Zweifelsfällen als Entscheidungskriterium jedoch nur inhaltliche Zusammengehörigkeit sowie Vorkommen von Häsitationspartikeln und Pausen herangezogen werden.

Selbstverständlich ist auch denkbar, daß eine äußerungsinterne freie Phrase weder dem vorausgehenden noch dem nachfolgenden Satz zuzuordnen ist, sondern als vollkommen eigenständige Phrase einzuordnen ist, vergleichbar mit der freien Phrase ohne Matrixsatz. Entsprechendes gilt für freie Phrasen links bzw. rechts eines Matrixsatzes. Um die Anzahl der Zweifelsfälle nicht zu erhöhen, wurde diese Kategorisierung nicht zugelassen.

Dieses Kategorisierungsproblem betrifft an sich einen wesentlichen Aspekt linguistischer Theoriebildung, nämlich die Bestimmung der Einheiten Satz, Äußerung etc. Eine solche Bestimmung konnte nicht Gegenstand vorliegender Untersuchung sein. Vielmehr wurde hier eine prätheoretische, durchaus praktikable Einteilung von häufig vorkommenden, aber theoretisch vernachlässigter Phänomene vorgenommen, die gerade zum Ziel hat, eine befriedigende Möglichkeit der Verarbeitung dieser Phänomene zu finden.

Unabhängig von ihrer Position ist in Tabelle 3 die kategorische bzw. strukturelle Ausprägung der freien Phrasen mit ihren Vorkommenshäufigkeiten zusammengestellt. Dabei sind einige der Ausprägungen (z.B. Linksversetzung, Rechtsversetzung, Nachtrag, Ausklammerung) auf eine bestimmte Position beschränkt.

Quantitativ gesehen spielen die integrierten freien Phrasen im Gegensatz zu den nicht–integrierten eine deutlich geringere Rolle.

Zu den satzintegrierten freien Phrasen werden hier links– und rechtsversetzte Strukturen, Nachträge und Ausklammerungen sowie neben– und bestimmte unterordnende Konjunktionen und Konjunktionaladverbien gezählt. Linksversetzte Strukturen sind satzgliedwertige Ausdrücke vor dem Vorfeld, die durch korreferente Proformen in den entsprechenden Satz integriert sind (vgl. z.B. (37–39)). Rechtsversetzte Strukturen sind satzgliedwertige Ausdrücke nach der Satzklammer, die durch korreferente Proformen in den entsprechenden Satz integriert sind (vgl. z.B. (40)). Sie sind im vorliegenden Korpus nicht belegt. Die Ausklammerung ist eine Nachfeldbesetzung durch satzgliedwertige Ausdrücke, ohne daß auf sie im dazugehörigen Satz mit Bezugselementen referiert wird (vgl. z.B. (41)). Ebenso wie bei einer Ausklammerung müssen auch die Ausdrücke eines Nachtrages als Positionsvariante in den Satzrahmen des entsprechenden Satzes einfügbar sein.

Tabelle3: **Ausprägungen der freien Phrase**

Klasse	%	Ø
integriert:		
Linksversetzung	1.40	.017
Rechtsversetzung	.00	.000
Nachtrag/Ausklammerung	4.52	.050
nebenordnende Konjunktion	5.46	.056
unterordnende Konjunktion	.00	.000
Konjunktionaladverb	7.49	.075
nicht–integriert:		
AdjP/AdvP	22.00	.262
NP	20.90	.242
PP	12.64	.181
VP	1.09	.011
Antwortpartikel, Antwortfloskel	31.36	.349
Gruß–/Dank–/Entschuldigung–Floskel	12.95	.161
andere Floskel	1.72	.018
Interjektion	5.93	.059

Auch sie haben im dazugehörigen Satz keine Proform als Bezugselement. Im Gegensatz zu Ausklammerungen können Nachträge als Satzellipsen verstanden werden, die durch bestimmte Floskeln, insbesondere *und zwar*, eingeleitet werden können (vgl. z.B. (42)). Da die Unterscheidung zwischen Ausklammerung und Nachtrag mitunter subtil ist oder sich nur in der Intonation äußert, wurden hier in der Auszählung diese zwei Strukturen zu einer einzigen zusammengefaßt.

(37) *den ein zwanzigsten April, den hätte ich Ihnen noch vorzuschlagen.*

(38) *wenn Sie nachmittags kommen könnten, das wäre okay.*

(39) *am Donnerstag, da habe ich schon einen Termin.*

(40) *#*[5] *er hat sie bestanden, die Lehre.*

(41) *wie sieht es mit Ihrer Bereitschaft aus am Wochenende zu arbeiten.*

(42) *könnten wir uns noch treffen, (und zwar) Mitte März irgendwann mal.*

Wie in geschriebener Sprache stehen neben– und unterordnenden Konjunktionen in gesprochener Sprache im Vorfeld. Im Unterschied zu geschriebener Sprache treten in gesprochener Sprache aber neben– und unterordnende Konjunktionen auf, ohne daß dazu innerhalb desselben Satzes oder im weiter vorausgehenden Diskurs ein erstes Konjunkt vorhanden ist (vgl. z.B. (43–45)). Im Gegensatz zu nebenordnenden Konjunktionen sind unterordnende Konjunktionen im Korpus nicht belegt (vgl. Tabelle 3).

(43) *und wie sieht's da so aus?*

(44) *dann aber bitte relativ früh.*

(45) *# weil/obwohl das ist besser.*

[5] Mit # werden Beispiele gekennzeichnet, die im Korpus nicht vorkommen.

Typisch für gesprochene Sprache sind Konjunktionaladverbien. Dies sind satzeinleitende Adverbien wie *also, so, nur, nun* etc., die in dieser Position mitunter nicht ihre Normalbedeutung haben (vgl. z.B. (46–47)). Sie kommen in 7,49 % der Äußerungen vor.

(46) *also der Montag sieht bei mir schlecht aus.*

(47) *so, jetzt bräuchten wir noch einen Termin.*

Freie Phrasen, die nicht durch ein Bezugselement oder eine spezifische Struktur wie Ausklammerung, Nachtrag oder Konjunktion etc. in einen Satz eingebunden sind, kommen wesentlich häufiger vor als integrierte freie Phrasen. Zu ihnen gehören AdjP's, AdvP's, NP's und PP's sowie spezifische Verbkonstruktionen, außerdem bestimmte Partikeln, Floskeln und Interjektionen (vgl. Tabelle 3).

Von den „Normalstrukturen“ kommen (unflektierte) AdjP's oder AdvP's (vgl. z.B. (48–50)) und NP's (vgl. z.B. (51–52)) am häufigsten vor, gefolgt von PP's (vgl. z.B. (53–54)).

(48) *sehr gut, prima, dann is' das ja kein Problem.*

(49) *wann? vierzehn Uhr is' wohl kaum angebracht.*

(50) *die zehnte Kalenderwoche, aber nur ungern.*

(51) *Moment, heute ist der zweite März.*

(52) *es geht nicht zwischen sieben und neun, keine Chance.*

(53) *zur Not ja, aber angenehmer wäre mir Februar.*

(54) *am sechzehnten kann ich nur bis sechzehn Uhr, nicht ab sechzehn Uhr.*

(55) *dann machen wir gleich ein' Termin fest, um elf Uhr.*

Relativ selten kommen dagegen Verbkonstruktionen vor. Dies sind gewöhnlich infinite Formen wie in (56–57). Vereinzelt sind auch finite Formen belegt (vgl. z.B. (58)).

(56) *mal sehen, also am dreizehnten is' ganz schlecht.*

(57) *nein, ausgeschlossen, bin ich außer Haus.*

(58) *ja, stimmt, 's war falsch.*

Die häufigste freie Phrase überhaupt ist eine Antwortpartikel oder –floskel wie *ja*, *doch*, *okay*, *alles klar*, *in Ordnung*. Sie kommt durchschnittlich 0,349 mal in jeder Äußerung des Korpus vor. Diese Zahl reflektiert gewiß die spezifische Domäne des Korpus und ist wahrscheinlich nicht repräsentativ für andere Domänen gesprochener Sprache. Auch Gruß–, Dankes– und Entschuldigungsfloskeln wie *hallo*, *auf Wiedersehen*, *bis dann*, *vielen Dank*, *sorry*, *tut mir leid* etc. kommen häufig vor, was durch die hohe Gesamtzahl an relativ kurzen Dialogen mitbedingt ist. Andere Floskeln, die aber selten vorkommen, sind z.B. *nun gut, das heißt, bitte*.

Die letzte Klasse der freien Phrasen sind Interjektion wie *tja*, *na ja*, *ach so*, *oh*, *ah*. Sie kommen beinahe ausschließlich in gesprochener Sprache vor und sind hier in 5,93 % der Äußerungen mindestesn 1 mal belegt.

5.2 Satzebene

In diesem Abschnitt werden Phänomene behandelt, die die Ebene des Satzes und nicht der Äußerung betreffen. Zunächst geht es um syntaktische Auffälligkeiten, danach um syntaktische Fehlleistungen.

Syntaktische Auffälligkeiten

Hier handelt es sich um genuin satzgrammatische Phänomene, die entweder typisch für gesprochene Sprache sind oder für die es auch in der Syntax zur geschriebenen Sprache keine „Standardbehandlung“ gibt. Von einer Ausnahme abgesehen, ist das Vorkommen solcher Phänomene, soweit sie überhaupt im Korpus belegt sind, recht marginal (vgl. Tabelle 4). Diese Ausnahme betrifft schwierige Adverbien und Partikeln bzw. Häufungen von Adverbien und/oder Partikeln (vgl. z.B. (59–61)). Sie kommen in jeder Äußerung durchschnittlich 0,385 mal vor. Ein ähnliches, aber nur selten vorkommendes Problem betrifft die Mehrfachbesetzung des Vorfelds durch Adverbialphrasen. Dies sind Fälle wie in (62–63)).

(59) *vormittags ginge es bei mir schon.*
(60) *dann wär' ja soweit alles klar.*
(61) *ich weiß halt eben noch nicht wie meine Planungen aussehen.*

(62) *und dann hinterher von halb zwei bis um vier könnten wir dann alles Weitere besprechen.*
(63) *so etwas bei einem Arbeits–Essen bespreche ich eigentlich ungerne.*

Von den verschiedenen Formen der Distanzstellung ist im Korpus nur eine belegt, nämlich die weite Distanzstellung wie in (64), und diese kommt auch nur in 0,31 % der Äußerungen vor. Andere mögliche Distanzstellungen wie die *was...für*–Konstruktion (z.B. 65)), Präpositionenstrandung (z.B. 66)), Distanzstellung von Determinierer und/oder Adjektiv (z.B. 67)) oder Distanzstellung eines NP–Attributs (z.B. 68)) sind im Korpus nicht belegt.

(64) *dann würde ich sagen treffen wir uns im Juni.*

(65) *# was gibt es denn da für Züge?*
(66) *# da kann ich nicht hin mitfahren.*
(67) *# Züge fahren keine.*
(68) *# von dem Haus sah man nur das Dach.*

Tabelle4: **Syntaktische Auffälligkeiten**

Klasse	%	Ø
Adverb/Partikel(häufung):		
nicht–triviale Adverb/Partikel(häufung)	31.98	.385
Adverbialhäufung im Vorfeld	1.09	.011
Distanzstellungen:		
Weite Distanzstellung	.31	.003
„was ... für“	.00	.000
Präpositionenstrandung	.00	.000
Distanzstellung von Determinierer und/oder Adjektiv	.00	.000
Distanzstellung von NP–Attribut	.00	.000
Ellipsen (außer freie Phrasen):		
eigentliche/uneigentliche Verbspitzenstellung	2.03	.020
elliptischer Matrixsatz mit konjunktionalem Nebensatz	1.40	.014
NP/PP/AdvP als Nebensatzellipse	3.43	.036
copula–freier konjunktionaler Nebensatz	.16	.002
freier Nebensatz (ohne Matrixsatz)	.47	.005
andere Auslassung	.00	.000
Parenthesen:		
parenthetische Floskel, Interjektion, Füllwort	.78	.009
parenthetischer Schaltsatz	.16	.002
parenthetischer Gliedsatz	.00	.000
ungleiche Koordinationen:		
Koordination von ungleichen Gliedern	3.12	.031
V_letzt–V2–Änderung in Satzkoordination	.00	.000
andere syntaktische Auffälligkeiten:		
Direkte statt indirekte Rede	3.59	.039
Aufforderung (V_erst)	2.34	.023
feste Redewendung oder Mehrwort–Ausdruck	4.37	.053
ungewöhnliche/schwierige Struktur	6.55	.069

Die Variation und Vorkommenshäufigkeit bei den Ellipsen ist etwas größer. Abgesehen von den freien Phrasen, die hier nicht als Ellipsen gezählt wurden, kommt die eigentliche und die uneigentliche Verbspitzenstellung als Vorfeldellipse vor. Während bei der eigentlichen Verbspitzenstellung eine expletives *es* oder *da, dann* ausgelassen wird (vgl. z.B. (69–70)), fehlt bei der uneigentlichen Verbspitzenstellung ein obligatorisches Argument wie z.B. in (71–72).

(69) *also [ ] würd' ich mal elf Uhr sagen.*
(70) *[ ] bin ich im Urlaub.*

(71) *[ ] ist richtig.*
(72) *[ ] war auch von mir 'n Fehler.*

Als weiterer Typ von Ellipse kommt ein elliptischer Matrixsatz mit konjunktionalem Nebensatz vor. Belege sind z.B. (73–74).

(73) *<u>gut</u>, daß ich Sie treffe.*
(74) *wenn ich noch Zeit zum Essen habe, <u>kein Problem</u>.*

Der am häufigsten belegte Ellipsentyp sind NP's, PP's oder AdvP's als Nebensatzellipsen wie in (75–77).

(75) *ich dachte <u>in der nächsten Woche</u>.*
(76) *es sei denn <u>abends</u>.*
(77) *sagen wir lieber <u>vierzehn Uhr</u>.*

Darüberhinaus gibt es marginale Vorkommenshäufigkeiten für konjunktionale Nebensätze, denen eine Kopula und das Subjekt fehlt (z.B. (78)), und für Nebensätze, die keinen Matrixsatz haben (z.B. (79)). Weitere Formen der Ellipse sind nicht belegt.

(78) *im Januar, wenn möglich.*

(79) *obwohl das auch nich' viel besser ist.*

Parenthesen treten im Korpus kaum auf. Von den wenigen Belegen sind die meisten floskelartige Ausdrücke wie (80–82). Außerdem sind Schaltsätze als Parenthesen belegt (z.B. (83)), Gliedsätze dagegen nicht.

(80) *wie würd's denn bei Ihnen da so um naja, sagen wir mal vierzehn Uhr aussehen?*

(81) *da hätt' ich glaub' ich die Freundlichkeit dazu.*

(82) *das is' mir ehrlich gesagt zu spät.*

(83) *wir können auf den sechzehnten März, das ist ein Dienstag, ausweichen.*

Koordinationen von ungleichen Gliedern wie in (84–86) kommen in ca. 3 % der Äußerungen vor. Eine Änderung der Verbstellung (V_letzt statt V2) im zweiten Teil einer Satzkoordination wie z.B. in (87) ist nicht belegt.

(84) *ich habe davon gehört und habe den Termin freigehalten.*

(85) *Sie fragen nach dem Botanischen Garten und kommen über den Fasanen–Garten rüber.*

(86) *wir müssen Zugeständnisse machen und am Wochenende uns treffen.*

(87) *# wenn ich bald ankomme und ich gehe dann, ...*

Andere im Korpus belegte syntaktische Auffälligkeiten, die typisch für gesprochene Sprache sind, sind der Gebrauch von direkter anstatt indirekter Rede (vgl. z.B. (88–89)) und Aufforderungen mit V_erst–Stellung (vgl. z.B. (90)). Letztere unterscheiden sich nur prosodisch von einer entsprechenden Frage.

(88) *ich würde vorschlagen, wir treffen uns um neun Uhr.*
(89) *meinen Sie, das reicht uns?*

(90) *lassen wir's dabei.*

Mit in die Auszählung mitaufgenommen wurden auch feste Redewendungen und Mehrwort–Äußerungen wie in (91–92). Sie kommen in 4,37 % der Äußerungen vor.

(91) *gehe ich da richtig in der Annahme?*
(92) *von mir aus können wir eine Stunde ansetzen.*

In einer Restkategorie wurden andere syntaktisch ungewöhnliche und/oder schwierige Konstruktionen wie in (93–97) aufgenommen. Sie kommen in 6,55 % der Äußerungen vor.

(93) *das mit dem schnell ist so 'ne Sache.*
(94) *bis fünfzehn Uhr müßte reichen.*
(95) *neuer Info-Bau zweiter Stock Zimmer zwei hundert drei.*
(96) *das Erscheinen um zum Aufwärmen ist unwahrscheinlich.*
(97) *ein zwei Stunden*

Syntaktische Fehlleistungen

Unter syntaktischen Fehlleistungen werden hier Phänomene verstanden, die man im allgemeinen auch in bezug auf gesprochene Sprache als agrammatisch beurteilt. Durchschnittlich kommen in jeder Äußerung 0,176 syntaktische Fehlleistungen vor. Hierzu gehören Neuansätze, die mitunter zusammen mit Häsitationspartikeln und/ oder Pausen auftreten (vgl. z.B. (98–100)).

(98) *ich würde sagen [wir]* *pause* *ein Termin am siebzehnten wäre nicht schlecht.*

(99) *vierzehn Uhr [muß ich wieder] hab' ich einen Kunden–Termin.*

(100) *wir könnten ein' Aperitif in meiner Strand–Villa [in] bei Monaco machen.*

Für eine detaillierte Analyse von Neuansätzen wäre zu unterscheiden, in welcher Position im Satz neu angesetzt wird (z.B. in einer Phrasengrenze, Wortgrenze oder wortintern), welche Einheit in dieser Position eingesetzt wird (z.B. Satz, Phrase, Wort oder Silbe) und ob gleichzeitig eine Häsitationspartikel und/oder Pause auftritt. Für eine solche Detaillanalyse gibt es im Korpus nicht genügend Belege. Es wurden deshalb alle Formen von Neuansätzen zu einer einzigen Klasse zusammengefaßt. In jeder Äußerung kommen durchschnittlich 0,08 Neuansätze vor (vgl. Tabelle 5). Unabhängig von einer Kookkurenz mit Neuansätzen wurde die Vorkommenshäufigkeit von Häsitationen und Pausen bereits oben in Tabelle 1 angeführt.

Tabelle5: **Syntaktische Fehlleistungen**

Klasse	%	Ø
Neuansatz	6.40	.080
Abbruch (technisch bedingt)	6.40	.064
Kongruenzfehler, Genusfehler	.94	.009
andere agrammatische Struktur	2.18	.023

Bei der Verwendung eines möglichen Verbmobil–Gerätes können agrammatische Strukturen nicht nur durch die Sprachperformanz des Sprechers, sondern schlicht durch eine fehlerhafte manuelle Betätigung des Gerätes, insbesondere durch zu frühen oder zu späten Tastendruck, bedingt sein (vgl.

z.B. (101)). Da dies in 6,4 % aller Äußerungen mindestens 1 mal vorkommt, sollte das Übersetzungssystem robust genug sein, auch solche agrammatischen Strukturen verarbeiten zu können.

(101) {cutoff} *prächtig an. ich glaube, das machen wir.*

Auffallend wenig dagegen kommen Kongruenz– und Genusfehler wie z.B. in (102–103) vor.

(102) *ich hab' im ersten März–Woche nichts Größeres vor.*

(103) *am darauffolgendem Freitag.*

Auch andere agrammatische Strukturen wie Anakoluth oder Kontamination kommen auffallend wenig vor (vgl. z.B. (104–105)), zumal zu dieser Klasse auch bestimmte Dialektformen wie in (106) gezählt wurden.

(104) *könnten wir's auf Montag probieren.*

(105) *das ganz einfach.*

(106) *'s wär' gut, wenn mir einen Termin ausmachet.*

6 Schlußfolgerungen

Die Art und Vorkommenshäufigkeit der Phänomene, wie sie im Rahmen dieser Untersuchung aufgezeigt wurden, lassen für die weiteren Verbmobil–Arbeiten in den Bereichen Grammatik und Parsing einige Schlußfolgerungen zu:

- Äußerungen, die die Restriktionen gängiger Sprachverarbeitungssysteme erfüllen, d.h. je Eingabe genau einen vollständigen Satz ohne freie Phrasen, kommen in der Domäne, in der ein Verbmobil–Gerät eingesetzt werden soll, relativ wenig vor (in 13 % der Äußerungen). Deshalb ist eine Satzgrammatik zwar notwendig, aber keinesfalls ausreichend.
- Für die Behandlung von Mehrsatz–Äußerungen und Sätzen mit freien Phrasen ist die Feststellung der Satzgrenzen das dominante Problem. Da diese Strukturen in mehr als 70 % aller Äußerungen vorkommen, muß deren Behandlung im Vordergrund stehen, wenn man ein lauffähiges System entwickeln will, das den in quantitativer Hinsicht vordringlichsten Anforderungen eines realen Einsatzes genügt.
- Einen wesentlichen Beitrag zur Lösung dieses Problems kann die Berücksichtigung prosodischer Information leisten.
- Phänomene, deren Behandlung die linguistische Theorie vorantreiben könnte, sind quantitativ marginal. Ihre Lösung trägt zur Effizienz eines einsatzfähigen Verbmobil–Gerätes nicht viel bei. Dagegen würde die Beherrschung „linguistisch unattraktiver" Phänomene wie z.B. Äußerungseinleitung, Adverb– und Partikelhäufung oder komplexe Datums– und Zeitangaben die Effizienz wesentlich steigern.
- *Feature relaxation*–Methoden können die Robustheit der Analyse nur unwesentlich verbessern, da Kongruenzfehler beim Sprechen kaum vorkommen.

- Wichtiger für die Robustheit sind stattdessen die Behandlung unvollständiger Sätze und die Behandlung von Neuansätzen.

Literatur

[Altmann 1981] Altmann, H. (1981). *Formen der »Herausstellung« im Deutschen. Rechtsversetzung, Linksversetzung, Freies Thema und verwandte Konstruktionen.* Niemeyer, Tübingen.

[Auer 1991] Auer, P. (1991). Zur Verbspitzenstellung im gesprochenen Deutsch. Vortrag DGfS–Jahrestagung 1991, Aachen. unveröffentlichtes Ms.

[Bayer 1973] Bayer, K. (1973). Verteilung und Funktion der sogenannten Parenthesen in Texten gesprochener Sprache. *Deutsche Sprache*, 1: 64–115.

[Betten 1976] Betten, A. (1976). Ellipsen, Anakoluthe und Parenthesen. Fälle für Grammatik, Stilistik, Sprechakttheorie oder Konversationsanalyse? *Deutsche Sprache*, 3: 207–230.

[Engel 1974] Engel, U. (1974). Syntaktische Besonderheiten der deutschen Alltagssprache. In *Gesprochene Sprache*, pages 199–228, Düsseldorf. Schwann.

[Höhne–Leska 1975] Höhne–Leska, C. (1975). *Statistische Untersuchungen zur Syntax gesprochener und geschriebener deutscher Gegenwartssprache.* Akademie–Verlag, Berlin.

[Horacek et al. 1988] Horacek, H., H. Bergmann, R. Block, M. Fliegner, M. Gerlach, M. Poesio, und M. Sprenger (1988). From meaning to meaning: A walk through WISBER's semantic–pragmatic processing. Technical Report 30, Universität Hamburg.

[Klein 1985] Klein, W. (1985). Ellipse, Fokusgliederung und thematischer Stand. In R. Meyer-Hermann und H. Rieser, editors, *Ellipsen und fragmentarische Ausdrücke*, pages 1–24, Tübingen. Niemeyer.

[Krause und Hitzenberger 1992] J. Krause und L. Hitzenberger, editors (1992). *Computer Talk*, Hildesheim. Georg Olms Verlag.

[Niedermair et al. 1990] Niedermair, G. T., M. Streit, und H. Tropf (1990). Linguistic processing related to speech understanding in SPICOS II. *Speech Communication*, 9: 565–585.

[Rath 1975] Rath, R. (1975). Korrektur und Anakoluth im Gesprochenen Deutsch. *Linguistische Berichte*, 37: 1–12.

[Schachtl und Block 1991] Schachtl, S. und H. U. Block (1991). Syntaktische Beschreibung in Systemen zur Verarbeitung gesprochener Sprache. Technical Report ASL-TR-11-91/SIM, Siemens AG, Zentralabteilung Forschung und Entwicklung, München.

[Schank und Schoenthal 1976] Schank, G. und G. Schoenthal (1976). *Gesprochene Sprache. Eine Einführung in Forschungsansätze und Analysemethoden.* Niemeyer, Tübingen.

[Steger 1987] Steger, H. (1987). Bilden „gesprochene Sprache“ und „geschriebene Sprache“ eigene Sprachvarietäten? In H. Aust, editor, *Festgabe für Theodor Lewandowsky zum 60. Geburtstag.*, pages 35–58, Tübingen. Narr.