

# Annotations and Tools for an Activity Based Spoken Language Corpus

**Jens Allwood, Leif Grönqvist,  
Elisabeth Ahlsén and Magnus Gunnarsson**

Dep. of Linguistics, Göteborgs University  
Box 200

SE-405 30 Göteborg, SWEDEN

{jens,leifg,eliza,mgunnar}@ling.gu.se

## Introduction

The paper contains a description of the Spoken Language Corpus of Swedish at the Department of Linguistics, Göteborg University (GSLC), and a summary of the various types of analysis and tools that have been developed for work on this corpus. Work on the corpus was started in the late 1970:s. It is incrementally growing and presently consists of 1.3 million words from about 25 different social activities. The corpus was initiated to meet a growing interest in naturalistic spoken language data. It is based on the fact that spoken language varies considerably in different social activities with regard to pronunciation, vocabulary, grammar and communicative functions. The goal of the corpus is to include spoken language from as many social activities as possible to get a more complete understanding of the role of language and communication in human social life. This type of spoken language corpus is still fairly unique even for English, since many spoken language corpora (certainly for Swedish) have been collected for special purposes, like speech recognition, phonetics, dialectal variation or interaction with a computerized dialog system in a very narrow domain, e.g. (Map Task (Isard and Carletta (1995), TRAINS (Heeman and Allen 1994), Waxholm (Blomerg et al. 1993). Compared to English corpora, the Göteborg corpus is most similar to the Wellington Corpus of Spoken New Zealand English (Holmes, Vine and Johnson 1998), but also has traits in common with the BNC, the London/Lund corpus (Svartvik

1990) and the Danish BySoc corpus (Gregersen 1991, Henrichsen 1997).

The corpus is based on audio (50%) or video/audio (50%) recordings of naturalistically occurring interactions.

The recordings have been transcribed according to a transcription standard consisting of a language neutral part – presently Göteborg transcription standard, GTS 6.2 (Nivre 1999a) (it has been tested on Chinese, Arabic, English, Spanish, Bulgarian and Finnish) and a language particular part concerned with Swedish – presently Modified Standard Orthography, MSO6 (Nivre 1999b). Both parts have undergone 6 major revisions and several minor ones. In order to enhance the reliability, all transcriptions are manually checked by another person than the transcriber. They are also checked for correctness of format, before they are inserted into the corpus. In MSO, standard orthography is used unless there are several spoken language pronunciation variants of a word. When there are several variants, these are kept apart graphically. Although the goal is to keep transcription simple, the standard includes features of spoken language such as contrastive stress, overlaps and pauses. It also includes procedures for anonymizing transcriptions and for introducing comments on part of the transcription.

Below, we will also describe several tools we have developed for using the corpus. The tools have, like the corpus, been incrementally developed since the early 1980:s and are all

concerned with work on the corpus. Using the tools and the corpus, we have done various kinds of quantitative and qualitative analysis, an example of this is a book of frequencies of Swedish spoken language. The book contains word frequencies both for the words in MSO format and in standard orthographic format. It also contains comparisons between word frequencies in spoken and written language (cf. Allwood 1998). There is statistics on the parts of speech represented in the corpus, based on an automatic probabilistic tagging, yielding a 97% correct classification. In addition, the corpus has been the basis for work using various kinds of manual coding, e.g. communication management (including hesitations, changes, feedback and turntaking), speech acts, obligations, misunderstandings, etc (cf. Allwood 2001). The corpus can also be used for other types of qualitative analysis, e.g. for CA-related sequential analysis. The recordings in the corpus are continuously being digitized on to DV tapes and/or CD:s with Mpeg compression. Each CD contains both transcriptions and recordings.

### 1. GSLC and other corpora

The spoken language corpora, besides GSLC include several other corpora, cf. table 1 below. We also work with other spoken language corpora collected by other teams.

**Table 1.** Spoken language corpora at Göteborg University, Department of Linguistics (Parts of the corpora are based on multimodal recordings.)

- Göteborg Spoken Language Corpus – GSLC (Kernel Corpus - adult first language Swedish), 1.3 million words
- Adult language learners of Swedish, 2 million words
- Aphasic speakers
- Child language corpus (Swedish and Scandinavian), 0.75 million words including the adults
- Educational progress, 416 interviews, 2 million words
- Non-Swedish adult spoken language corpus
  - \* Chinese (70 000 words)
  - \* Bulgarian (25 000 words)
  - \* Arabic
  - \* English (10 000 words), BNC
  - \* Finnish
  - \* Italian (3000 words)
  - \* Norwegian (140 000 words)
  - \* Spanish
- Wizard-Of-Oz Corpus, Bionic
- Intercultural communication corpus

GSLC, the kernel corpus of adult first language Swedish speakers is the corpus we will focus on in this article. In Table 2, below, we present basic data on this corpus. However, regroupings of, or selections from, the corpus according to such criteria are also possible. The limitations which exist in our ability to create subcorpora are dependent on the fact that we do not always have the relevant information about individual speakers.

**Table 2.** The Göteborg Spoken Language Corpus

Type of social activity	No. of recordings	Average number of speakers / recording	No. of sections *	Tokens (including pauses and comments)	Audible word tokens uttered in recording	Duration**
Auction	2	6.0	111	26 776	26 459	3:14:11
Bus driver/ passenger	1	33.0	20	1 360	1 345	0:13:33
Consultation	16	3.0	239	34 865	34 285	2:44:25
Court	6	5.0	79	33 401	33 261	3:58:33
Dinner	5	8.0	30	30 738	30 001	2:49:54

Discussion	34	5.8	255	240 426	237 583	17:19:24
Factory conversation	5	7.4	48	29 024	28 860	2:19:47
Formal meeting	13	9.7	186	219 352	215 582	15:45:54
Hotel	9	19.2	183	18 950	18 137	6:47:50
Informal conversation	22	4.4	152	94 490	93 436	7:48:41
Information Service	32	2.1	40	14 700	14 614	0:13:40
Interview	58	2.9	1 031	396 758	393 907	30:34:27
Lecture	2	3.5	3	14 682	14 667	1:38:00
Market	4	24.2	38	12 581	12 175	2:18:37
Religious Service	2	3.5	10	10 273	10 234	1:10:45
Retelling of article	7	2.0	7	5 331	5 290	0:42:00
Role play	2	2.5	7	5 702	5 652	0:39:16
Shop	49	7.4	139	36 385	34 976	6:40:46
Task-oriented dialog	26	2.3	46	15 475	15 347	2:05:20
Therapy	2	7.0	8	13 841	13 529	2:04:07
Trade fair	16	2.1	16	14 353	14 116	1:12:46
Travel agency	40	2.7	112	40 370	40 129	5:53:57
<b>Total</b>	<b>353</b>	<b>4.9</b>	<b>2 762</b>	<b>1 310 284</b>	<b>1 204 029</b>	<b>118:15:53</b>

\*A section is a longer phase of an activity with a distinct subordinate purpose. The bus driver/-passenger recording, for example, has 20 sections, where each section involves talk with a new passenger.

\*\* For some recordings, there is no duration available. We estimate that the figures given above probably under-represents actual duration by about 30 hours.

## 2. Storage

Around 50% of our 1.3 million tokens corpus is stored on audio tapes and the rest on video tapes (Umatic, VHS or BetaCAM). In order to preserve the recordings, analog tapes are being digitized onto DV tapes and/or Cd:s using Mpeg compression. One mini DV-tape holds 60 minutes and a DVCam 180 minutes. This format requires a fast computer. Using Mpeg compression, we have tried to use a constant data rate of around 200 kb per second. This

will give a fair quality and the format may be used on almost any PC/Mac.

## 3. Description of the corpus transcription standard

The transcription standard we have used (GTS + MSO) can perhaps most rapidly be explained through exemplification.

**Example 1.** Transcription according to the GTS + MSO standard with translation.

§1. Small talk

\$D: säger du de{t} ä{r} de{t} ä{r} de{t} så \$D: oh I see is it it is so troublesome then  
besvärlit då

\$P: ja ja \$P: yes yes

\$D: m // ha / de{t} kan ju bli så se{r} du \$D: m // yes / it can be that way you see

\$P: < jaha > \$P < yes >

@ <ingressive> @ <ingressive >

\$D: du ta{r} den på morronen \$D: you take it in the morning

\$P: nej inte på MORRONEN kan ja{g} ju tar \$P: no not in the MORNING I always take a  
allti en promenad på förmiddan [1 å0 ]1 då walk before lunch [1 and ]1 then I don't want  
vill ja{g} inte ha [2 den ]2 medicinen å0 sen [2 that ]2 medicine and then when I get home  
nä ja{g} kommer hem möjligtvis possibly

\$D: [1 {j}a ]1 \$D: [1 yes ]1

\$D: [2 nä ]2 \$D: [2 no ]2

The example shows the following properties of the transcription standard:

- (i) Section boundaries paragraph sign (§). These divide a longer activity up into subactivities. A doctor-patient interview can, for example have the following subactivities. (i) greetings and introduction, (ii) reason for visit, (iii) investigation, (iv) diagnosis, (v) prescribing treatment.
- (ii) Words and spaces between words.
- (iii) Dollar sign (\$) followed by capital letter, followed by colon (:) to indicate a new speaker and a new utterance.
- (iv) Double slash (//) to indicate pauses. Slashes /, // or /// are used to indicate pauses of different length.
- (v) Capital letters to indicate contrastive stress.
- (vi) Word indices to indicate which written language word corresponds to the spoken form given in the transcription (å0 corresponds to written language *och*). In the cases where spoken language variants can be viewed as abbreviated forms of written language, we use curly brackets {} to indicate what the standard orthographic form would be, e.g. de {t} = *det*.
- (vii) Overlaps are indicated using square brackets ([ ]) with indices which allow disambiguation if several speakers overlap simultaneously.
- (viii) Comments can be inserted using angular brackets (< >) to mark the scope of the comment and @< > for inserting the actual comment. These comments are about events which are important for the interaction or about such things as voice quality and gestures.

#### 4. Tools which have been developed

The following tools have been developed to aid work related to the corpora.

##### 4.1 TransTool

TransTool (Nivre et al., 1998) is a computer tool for transcribing spoken language in accordance with the transcription standard

(Nivre 1999a and b). It will help the user to transcribe correctly and make it much easier to keep track of indices for overlaps and comments (cf. Nivre et al. 1998).

##### 4.2 The Corpus Browser

The Corpus Browser is a web interface that makes it possible to search for words, word combinations and phrases (as regular expressions) in the Göteborg Spoken Language Corpus. The results can be presented as concordances or lists of expressions with as much context as you wish and with direct links to the transcription.

##### 4.3 TRACTOR

Tractor is a coding tool which makes it possible to create new coding schemas and annotate transcriptions. Coded segments can be discontinuous and it is also possible to code relations. A coding schema can be represented as a tree with strings on all nodes and leaves, and a coding value is a path through the tree. That model is similar to the file and folder structure on a computer harddisk. This framework makes it easy to analyze the codings in a Prolog system, but it is not possible to order the codings or code a coding, because a coding only consists of two discontinuous intervals and a coded value (Larsson 1997).

##### 4.4 Visualization of codings with FrameMaker

We have also created a toolbox that makes it possible to visualize coding schemas and coding values with colors, bold, italics, etc. directly in the transcription as a FrameMaker document. Different parts of the transcription may also be marked (or be excluded) to get a legible view of it without details you might not be interested in for the moment (Grönqvist 1999).

##### 4.5 TraSA

If you have a corpus transcribed according to the Göteborg Transcription Standard, using TraSA it is very easy to calculate some 30 statistical measurements for different properties, activities, sections and/or speakers

(Grönqvist 2000b). For example, you will be able to count things like number of tokens, types, utterances, or more complex things like, theoretical vocabulary.

#### 4.6 SyncTool

SyncTool (Nivre et al., 1998) is developed (as a prototype for MultiTool) for synchronizing transcriptions with digitized audio/video recordings. It is also meant to be a viewing tool allowing the researcher to view the transcription and play the related recording without having to manually locate the specific passage in the recording.

#### 4.7 Work on a synchronizing tool – MultiTool

Many of the tools above would be more useful if you could use their functionality simultaneously in one tool. MultiTool is an attempt to build such a general tool for linguistic annotation and transcribing of dialogs, as well as for browsing, searching and counting. The system can handle any number of participants, overlapping speech, hierarchical coding schemas, discontinuous coding intervals, relations, and synchronization between codings and the media files (see Grönqvist 2000a).

The fundamental idea is to collect all information in an internal state which is a low level representation of all kinds of annotations, including the transcription, containing the abstract objects: codings and synchronizations. These are the basic types of information the computer program requires. For researchers using the audio/video recordings of the corpus, the transcriptions are merely a coding of the recordings. One important detail is that views (e.g. “partiture” and other views of transcription, views of codings, acoustic analysis as well as audio and video files) pertaining to the same point in time can be synchronized to show the same sequence from different points of view whenever the user scrolls only in one of them. The internal state contains all the information so it is possible to have many different views of the same sequence of the dialog. Changes made in one view will immediately change the internal state and as a consequence the other views.

MultiTool is written in JAVA+JMF which makes it fairly platform-independent and since interpreters are rapidly getting more efficient, the performance will probably be good enough on the major platforms very soon. One main feature we will add is the import and export functions for our local transcription format, TRACTOR files and probably also for the CA (“conversation analysis”) format.

For many users, the newer versions of MultiTool will in the future replace all the tools above. However, TraSA and the Corpus Browser will still be needed when working on bigger portions of the corpus at the same time. With the appropriate import/export functions different users will be able to use their own transcription- and annotation formats with MultiTool. In our opinion the features in MultiTool will be a good base level for things to do with a multimodal spoken language corpus: transcribing, coding/annotating, converting, searching, counting, browsing, visualizing. For some other user profiles there are better tools, like Waves for phoneticians, and MediaTagger for simpler annotations.

### 5. Types of quantitative analysis

Using the information provided by the transcriptions following the Göteborg standard, we have defined a set of automatically derivable properties which include the following (cf. Allwood and Hagman 1994, Allwood 1996):

- (i) **Volume:** Volume comprises measures of the number of words, word length, pauses, stresses, overlaps, utterances and turns relative to speaker, activity and subactivity.
- (ii) **Ratios:** Various ratios can then be calculated based on the volume measures.

For example:

$$\begin{aligned} \text{MLU} &= \text{words/utterances} \\ \% \text{ pauses} &= 100 * \text{pauses}/(\text{words} + \text{pauses}) \\ \% \text{ stress} &= 100 * \text{stressed words}/\text{words} \\ \% \text{ overlap} &= 100 * \text{overlapped words}/\text{words} \\ \text{speed} &= \text{words}/\text{duration} \end{aligned}$$

Alternatively, pause, stress and overlap can be given per utterance. All of these measures can then be relativized to speaker, activity or subactivity.

- (iii) **Special descriptors:** One example of a special type of descriptor is “vocabulary richness” as measured through type/token, Guiraud, Über, Herdan or “theoretical vocabulary”, cf. van Hout & Rietveld (1993). Other descriptors we have constructed are “stereotypicality” which looks at how often words and phrases are repeated in an activity, “verbal dominance” and “verbal equality”, “liveliness” and “caution”, and overlap in different utterance positions.
- (iv) **Lemma:** We also implemented a simple stemming algorithm which enables us to collect regularly inflected forms together with their stem.
- (v) **Parts of speech:** Parts of speech are assigned using a probability based statistical (Viterbi - trigram) parts of speech tagger which has been adapted to spoken language. Using this, a parts of speech coding has been done for the whole Göteborg Spoken Language Corpus, roughly 1.3 million transcribed words. The correctness of the coding is about 97% (cf. Nivre & Grönqvist, 2001). Words subdivided according to parts of speech can then be assigned to speaker, activity or subactivity.
- (vi) **Collocations:** All speakers, activities and subactivities can be characterized in terms of their collocations, sorted by frequency as complete utterances or by “mutual information” (Manning and Schütze 1999).
- (vii) **Frequency lists:** Frequency lists can be made for words, lemmas, parts of speech, collocations, and utterance types.
- (viii) **Sequences of parts of speech:** Utterances of different length can be characterized as to sequence of parts of speech. This allows a first analysis of grammatical differences between speakers, activities and subactivities.
- (ix) **Similarities:** Similarities between activities are captured by looking at the extent to which words and collocations are shared between activities.

Validity and reliability checks are done manually for all automatic measures.

## 6. Types of qualitative analysis

### 6.1 Overview

In order to increase the reliability, qualitative analysis in Göteborg has often resulted in the development of coding schemas, by which we mean schemas for annotations on top of the transcription. If the Göteborg coding is compared to other coding schemas, we can see that some lie on top of transcription, e.g. DAMSL (Core and Allen, 1997) and DRI, while others are being integrated with the transcription standard, e.g. the MATE markup framework (Dybkjær et al.1998). A fair comparison between the major, not to mention all, schemas is beyond the scope of this paper. The coding schemas presented here reflect the areas of interest that the Göteborg group have focussed on. The underlying transcription standard naturally restricts the level of granularity for any new coding schemas, but the two coding tools developed in Göteborg, MultiTool and TRACTOR, are meant to be as independent of any individual coding schema or transcription standard as possible. The following list provides an overview of the Göteborg coding schemas (cf. Allwood 2001):

1. Social activity and Communicative act related coding
  - 1.1 Social activity
  - 1.2 Communicative acts
  - 1.3 Expressive and Evocative functions
  - 1.4 Obligations
2. Communication management related coding
  - 2.1 Feedback
  - 2.2 Turn and sequence management
  - 2.3 Own Communication Management
3. Grammatical coding
  - 3.1 Parts of speech (automatic, probabilistic)
  - 3.2 Maximal grammatical units
4. Semantic coding.

Reliability checking is planned to be included in the development of all coding schemata. So far, the coding of Feedback and Own Communication Management has been checked for inter-rater reliability (using Cohen’s kappa).

## 6.2 Contributions, utterances and turns

Following Grice (1975), Allwood, Nivre and Ahlsén (1990) and Allwood (2000), the basic units of dialog are gestural or vocal contributions from the participants. The term contribution is used instead of utterance in order to cover also gestural and written input to communication. Verbal contributions can consist of single morphemes or be several sentences long. The term turn is used to refer to the right to contribute, rather than to the contribution produced during that turn. One may make a contribution without having a turn and one may have the turn without using it for an active contribution, as demonstrated in the example below, in which B's first contribution involves giving positive feedback without having the turn (square brackets indicate overlap) and his/her second contribution involves being silent and doing nothing while having the turn.

A: look ice cream [would] you like an ice cream  
B1: [yeah]  
B2: (silence and no action)

Contributions, utterances and turns are not coded since they are obtainable directly from the Göteborg transcription standard.

## 6.3 Coding related to Social activity and Communicative acts

### 6.3.1 Social activity

Each transcription is linked to a database entry and a header containing information on:

- (i) The purpose, function and procedures of the activity
- (ii) The roles of the activity
- (iii) The artefacts, i.e. objects, furniture, instruments and media of the activity
- (iv) The social and physical environment
- (v) Anonymous categorical data on the participants, such as age, gender, dialect and ethnicity.

In addition, the major subactivities of each activity are given.

### 6.3.2 Communicative Acts

Each contribution can be coded with respect to one or more communicative acts which can occur sequentially or simultaneously. The communicative acts make up an extendible list, where often used types have been provided with definitions and operationalizations. Some often used types are the following: Request, Statement, Hesitation, Question, Answer, Specification, Confirmation, Ending interaction, Interruption, Affirmation, Conclusion, Offer.

### 6.3.3 Expressive and evocative functions

In accordance with Allwood (1976, 1978, 2000), each contribution is viewed as having both an expressive and an evocative function. These functions make explicit some of the features implied by the communicative act coding. The expressive function lets the sender express beliefs and other cognitive attitudes and emotions. What is "expressed" is made up of a combination of reactions to the preceding contribution(s) and novel initiatives. The evocative function is the reaction the sender intends to call forth in the hearer. Thus, the evocative function of a statement normally is to evoke a belief in the hearer, the evocative function of a question is to evoke an answer, and the evocative function of a request to evoke a desired action.

### 6.3.4 Obligations

If the dialog and communication is to be cooperatively pursued, whether it be in the service of some activity or not, they impose certain obligations on both sender and recipient. With regard to both expressive and evocative functions, the sender should take the receiver's perceptual, cognitive and behavioral ability into consideration and should not mislead, hurt or unnecessarily restrict the freedom of the receiver. The receiver should reciprocate with an evaluation of whether he/she can hear, understand and carry out the sender's evocative intentions and signal this to the interlocutor. The sender's and receiver's obligations can be summarized as follows (see also Allwood 1994):

**Sender:** 1. *Sincerity*, 2. *Motivation*,  
3. *Consideration* (cf. Allwood 1976)  
**Receiver:** 1. *Evaluation*, 2. *Report*,  
3. *Action*.

## 6.4 Communication management related coding

### 6.4.1 Introduction

The term “communicative management” refers to means whereby speakers can regulate interaction or their own communication. There are three coding schemas related to communication management (cf. Allwood 2001): 1) Feedback coding, 2) Turn and sequence management coding, and 3) Own Communication Management (OCM) coding.

### 6.4.2 Feedback coding schema

A feedback unit can be described as “a maximal continuous stretch of utterance (occurring on its own or as part of a larger utterance), the primary function of which is to give and/or elicit feedback concerning contact, perception, understanding and acceptance of evocative function” (Allwood, 1993). All feedback units are coded with respect to “Structure”, “Position/Status” and “Function”. Coding structure means coding grammatical category (part of speech, phrase or sentence) and also “structural operations”. “Structural operations” is subdivided into “phonological”, “morphological” and “contextual” operations, each of which have different values.

### 6.4.3 Turn and sequence management coding

Turn and sequence management coding encompasses the following phenomena:

(A) Overlap and interruption: Overlap is coded in the transcriptions and can be extracted automatically. Interruption is a code for those overlaps which aims at or succeed in changing the topic or taking away the floor from another speaker.

(B) Intended recipient: This type of coding has four self explanatory values

- (i) particular participant
- (ii) particular group of participants
- (iii) all participants
- (iv) no other participant (talking to

oneself).

(C) Marking of the opening and closing of subactivities and/or the interaction as a whole.

### 6.4.4 OCM coding schema

OCM means “Own Communication Management” and stands for processes that speakers use to regulate their own contributions to communicative interaction. OCM function coding concerns classifying whether the OCM unit is:

- choice related - helps the speaker gain time for processes concerning continuing choice of content and types of structural expressions, or:
- change related - helps the speaker change already produced content, structure or expression.

OCM units are also coded with respect to structure of the OCM related expression. This structure can be divided into “basic OCM features”, “basic OCM operations” and “complex OCM operations”. Pauses, simple OCM expressions such as hesitation sounds etc and explicit OCM phrases count as basic OCM features. Basic OCM operations are: “lengthening of continuants”, “self interruption” and “self repetition”. The category “Complex OCM operations” stands for different ways to modify the linguistic structure. The OCM coding schema is described in Allwood, Ahlsén, Nivre & Larsson (1997).

## 6.5 Grammatical coding

There are also ways of coding grammatical structure. One of these is the automatic coding of parts of speech mentioned above. Another is a coding of “The Maximal Grammatical Units”, a coding schema is described in Allwood (2001). When coding Maximal Grammatical Units, one should primarily try to find as large units as possible, the largest unit being complete sentences. Sentences are subclassified by using the schema “**sentences**”. In spoken language, there are many utterances that are not sentences, so secondarily, one should try to find complete phrases, which should be coded in the schema “**phrases**”. If it isn't possible to find either complete sentences or complete phrases, single words should be coded by parts of speech in the schema “Parts of speech”.

## 7. Conclusions and Future Directions

In this paper, we have described work done at the Department of Linguistics, Göteborg University to collect, transcribe and store spoken language material. We have also described some of the tools we have developed in order to aid work on analyzing the data both automatically and manually. Finally, we have described some of the results obtained so far. Future work will include incremental expansion of the corpus both to obtain data from new social activities and in order to equalize the size of the material from different activity types. We will also be making increased efforts to make the corpus more multimodal by making the audio and video recordings on which the transcriptions are based more available. Work on tools for analyzing the corpus will continue. The most immediate goal is to complete MultiTool which will hopefully give us a better possibility of working with multimodal data. Similarly, work on qualitative and quantitative analysis will be continued. An ambitious goal is to work toward a grammatical description of spoken language and toward a systematic description (perhaps not a grammar) of multimodal face-to-face communication.

## 8. References

- Jens Allwood (1976) *Linguistic Communication as Action and Cooperation*. "Gothenburg Monographs in Linguistics" 2. Göteborg University, Department of Linguistics, 257 p.
- Jens Allwood (1978) *On the Analysis of Communicative Action*. In "The Structure of Action", M. Brenner, ed., Basil Blackwell, Oxford, pp. 168-191.
- Jens Allwood (1993) *Feedback in Second Language Acquisition*, In "Adult Language Acquisition. Cross Linguistic Perspectives", Vol. II. C. Perdue, ed., Cambridge: Cambridge University Press, Cambridge, pp. 37-51.
- Jens Allwood (1994) *Obligations and Options in Dialogue*, Think, Vol 3, May, ITK, Tilburg University, 9-18.
- Jens Allwood, ed, (1996 and later editions) *Talspråksfrekvenser*, Ny och utvidgad upplaga. Gothenburg Papers in Theoretical Linguistics S21. Göteborg University, Department of Linguistics, 418 p.
- Jens Allwood (1998) Some Frequency based Differences between Spoken and Written Swedish. In Timo Haukioja, ed., *Proceedings of the 16th Scandinavian Conference of Linguistics*, Turku University, Department of Linguistics, pp. 18-29.
- Jens Allwood, (2000) *An Activity Based Approach to Pragmatics*. In "Abduction, Belief and Context in Dialogue; Studies in Computational Pragmatics", H. Bunt, & B. Black, eds., John Benjamins, Amsterdam, pp. 47-80.
- Jens Allwood, ed., (2001) *Dialog Coding – Function and Grammar: Göteborg Coding Schemas*. Gothenburg Papers in Theoretical Linguistics GPTL 85. Göteborg University, Department of Linguistics, 67 p.
- Jens Allwood and Johan Hagman (1994) *Some Simple Measures of Spoken Interaction*. In F. Gregersen, & J. Allwood, eds., "Spoken Language, Proceedings of the XIV Conference of Scandinavian Linguistics", pp. 3-22.
- Jens Allwood, Elisabeth Ahlsén, Joakim Nivre and Staffan Larsson (2001) *Own communication management*. In J. Allwood, ed., (2001) *Dialog Coding – Function and Grammar: Göteborg Coding Schemas*. Gothenburg Papers in Theoretical Linguistics GPTL 85. Göteborg University, Department of Linguistics, pp. 45-52.
- Jens Allwood, Joakim Nivre and Elisabeth Ahlsén (1990) *Speech Management: On the Non-Written Life of Speech*. Nordic Journal of Linguistics, 13, 3-48.
- Mats Blomberg, Rolf Carlson, Kjell Elenius, Björn Granström, Jonatan Gustafson, Sheri Hunnicutt, Roger Lindell and Lennart Neovius (1993) *An experimental dialogue system: WAXHOLM*, "Proceedings of EUROSPEECH 93", pp 1867-1870.
- Mark G. Core and James, F. Allen (1997) *Coding Dialogs with the DAMSL Annotation Scheme*. In Working Notes of AAAI Fall Symposium on Communicative Action in Humans and Machines, Boston, MA, November 1997.
- Laila Dybkjær, Niels Ole Bernsen, Hans Dybkjær, David McKelvie and Andreas Mengel (1998) *The MATE Markup Framework*. MATE Deliverable D1.2, November 1998, 15 p.
- Frans Gregersen (1991) *The Copenhagen Study in Urban Sociolinguistics*, 1+2; Reitzel, Copenhagen.
- H. Paul Grice (1975). *Logic and conversation*. In "Syntax and Semantics" Vol. 3: Speech Acts, P. Cole and J. L. Morgan, eds., Seminar Press, New York, pp. 41-58.

- Leif Grönqvist (1999) *Kodningsvisualisering med Framemaker*. Göteborg University, Department of Linguistics, 8 p.
- Leif Grönqvist (2000a) *The MultiTool User's Manual. A tool for browsing and synchronizing transcribed dialogues and corresponding video recordings*. Göteborg University, Department of Linguistics, 6 p.
- Leif Grönqvist (2000b) *The TraSA v0.8 Users Manual. A user friendly graphical tool for automatic transcription statistics*. Göteborg University, Department of Linguistics, 8 p.
- Peter A. Heeman and James, F. Allen (1994) *The TRAINS 93 Dialogues*. TRAINS Technical Note 94-2.
- Peter Juel Henriksen (1997) *Talesprog med Ansigtssløftning*, IAAS, Univ. of Copenhagen, Instrumentalis 10/97 (in Danish), 66 p.
- Janet Holmes, Bernadette Vine and Gary Johnson (1998) *Guide to the Wellington Corpus of Spoken New Zealand English*. Victoria University of Wellington, Wellington.
- Amy Isard and Jean Carletta (1995) *Transaction and action coding in the Map Task Corpus*. Research Paper HCRC/RP-65, 27 p.
- Staffan Larsson (1997) *TRACTOR v1.0b1 användarmanual*. Göteborg University, Department of Linguistics, 10 p.
- Christopher D. Manning and Hinrich Schütze (1999) *Foundations of Statistical Natural Language Processing*, The MIT Press, Boston, Mass., 620p.
- Joakim Nivre (1999a) *Transcription Standard. Version 6.2*. Göteborg University. Department of Linguistics, 38 p.
- Joakim Nivre (1999b) *Modifierad Standard-Ortografi (MSO) Version 6*. Göteborg University, Department of Linguistics, 9 p.
- Joakim Nivre, Kristina Tullgren, Jens Allwood, Elisabeth Ahlsén, Jenny Holm, Leif Grönqvist, Dario Lopez-Kästen and Sylvana Sofkova (1998) *Towards multimodal spoken language corpora: TransTool and SyncTool*. Proceedings of ACL-COLING 1998, June 1998.
- Joakim Nivre and Leif Grönqvist (2001) *Tagging a corpus of Spoken Swedish*. Forthcoming in International Journal of Corpus Linguistics.
- Roeland van Hout and Toni Rietveld (1993) *Statistical Techniques for the Study of Language and Language Behaviour*. Berlin & New York: Mouton de Gruyter, 400 p.
- Jan Svartvik (ed.) (1990), *The London Corpus of Spoken English: Description and Research*. "Lund Studies in English" 82. Lund University Press, 350 p.