A Cross-Linguistic Look at Discourse Connectives

Tatjana Scheffler

July 15, 2005

1   Introduction

Discourse connectives are parts of speech that encode relations between propositions, events, or situations. Connectives are usually seen as *links* between parts of discourse, which requires them to express semantic two-place relations. Thus, an adverb like *fortunately* does not constitute a discourse connective, because it takes only one semantic argument. However, the adverbial phrase *in addition* relates two propositions and thus fulfills the requirements for discourse connectives:

(1) Tom's car broke down. *Fortunately*, he was near a garage when it happened.

(2) Tom's car broke down. *In addition*, his cellphone didn't work when he tried to call the AAA.

   In this paper, I look at discourse connectives in German, comparing them to connectives in English. Syntactic properties will be in the center of my attention, because there are some interesting differences between the two languages with regard to word order and the classification of connectives. However, both the syntactic and semantic properties of discourse connectives will be discussed. It turns out that whereas some facts seem to be universal, others, for example what types of connectives exist (see section 2), are very language-specific.

2   Types of Discourse Connectives

2.1   English

Discourse Connectives in English fall into three classes, *coordinating conjunctions*, *subordinating conjunctions*, and *adverbials*. The three types are exemplified in (3–5), respectively.

(3) John likes apples, *but* Mary prefers pears.   (coord. conj.)

(4) John likes apples, *because* they are juicy.   (subord. conj.)

(5) John likes apples. *However,* these ones are sour.   (adverbial)

Syntactically, the conjunctions differ from adverbials significantly. Conjunctions take their arguments structurally, which means that the connective must be local to both its arguments. For coordinating conjunctions, this usually means that the connective appears between its first and second argument. Subordinating conjunctions in English can appear either sentence-initially, followed by both arguments (6), or medially, in between the two arguments (4).

(6) Because John likes fruit, he eats many apples per day.

Adverbials that are connectives have one structural argument. This is the clause in which they appear syntactically. The other argument can be non-local. This shows that the connective relates to this argument *anaphorically*, and not structurally (Webber et al., 2003). One can therefore say that adverbial connectives have two semantic arguments, but only one syntactic argument, because the other semantic argument is recovered anaphorically. More about the anaphoric argument of some connectives follows in section 3.

2.2 German

The syntax of German connectives has been extensively studied in Germany. A very detailed overview of German connectives can be found in (Pasch et al., 2003). Most of my analyses here are based on the data and descriptions there.

In German, the basic distinction between connectives with two syntactic arguments and those with one syntactic argument is also valid. Adverbials and particles have only one syntactic argument. At the same time, they are integrable (*integrierbar*) into their argument. Conjunctions, which take two syntactic arguments, cannot be integrated alone (without one of their argument sentences) into one of their argument clauses (*nicht konnektintegrierbar*).

(7) Das Meerschwein wurde gefangen, der Hase *aber* kam davon.[1]
    'The guinea pig was caught, *but* the rabbit got away.'

(8) Das Meerschwein wurde gefangen, der Hase (*und*) wurde freigelassen.
    int.: 'The guinea pig was caught, *and* the rabbit was released.'

Although this gives us the basic distinction between conjunctions and adverbials, the more fine-grained differentiation is more difficult than in English. The reason for this is the freer word order in German, which provides many positions for discourse connectives that are not available in English.

2.2.1 Conjunctions

Conjunctions do not divide neatly into coordinating and subordinating conjunctions. All conjunctions share the property that their arguments have to be realized locally to the connective. However, the relevant syntactic properties

---

[1] Note that *aber* (but) in German is not a coordinating conjunction, but rather shows a very specialized adverbial behavior. In this example, it is used to exemplify the positional possibilities of adverbials.

split the group of conjunctions into not just two, but four distinct classes. This is the case because *subordination* and *embedding* are distinct syntactic phenomena in German.

Subordination.

A syntactic item in a sentence *subordinates* its internal argument clause iff it requires verb-final word order in this clause.

Embedding.

A syntactic structure X is embedded in structure Y iff X is a constituent of Y that has a syntactic function in Y, but not vice versa.

That is, subordination can be detected by the presence of verb-final word order, whereas embedding is expressed by the ability of the embedded clause to occupy the Vorfeld of the embedding clause, or a position in its Mittelfeld, in addition to appearing in the Nachfeld.

The sentences (9–12) show examples of each of the combinations of [± Subordination] and [± Embedding].

(9) *Weil* er müde ist, geht er nach Hause.   [+ sub ; + emb]
'Because he is tired, he's going home.'

(10) Er wird schnell müde, *weshalb* er immer schon früh nach Hause geht.
'He gets tired easily, that's why he always goes home early.'

[+ sub ; − emb]

(11) *Vorausgesetzt*, er bekommt Urlaub, wollen wir nach Italien fahren.
'Provided that he gets vacation, we want to go to Italy.'

[− sub ; + emb]

(12) Der Chef ist im Urlaub *und* niemand kann ihn erreichen.
'The boss is on vacation and nobody can reach him.'   [− sub ; − emb]

German conjunctions can be broadly grouped into 4 classes that correspond to these feature matrices. Following Pasch et al. (2003), I distinguish subordinating conjunctions (Subjunktoren), which are both subordinating and embedding; postponers (Postponierer), which subordinate, but do not embed their internal argument; V2-embedders (Verbzweitsatzeinbetter), which embed (non-subordinated) clauses with verb-second word order; and the coordinating conjunctions, which are neither subordinating nor embedding. Table 1 summarizes the behavior of these classes. Some examples of German conjunctions are given in table 2.

2.2.2 Adverbials

The number of syntactic types of adverbial discourse connectives in German is very large, due to the distinct positions that a given adverbial can or cannot

	subord. conj.	postponers	V2-embedders	coord. conj.
subordinating?	yes	yes	no	no
embedding?	yes	no	yes	no
coordinating?	no	no	no	yes
int. word order	VF	VF	decl. V2	any
clause can appear in Vorfeld or Mittelfeld	yes	no	yes	no

Table 1: The four types of German conjunctions.

subord. conj.	postponers	V2-embedders	coordinating conj.
insofern, dass	als dass	angenommen	das heißt
nachdem	gdw.	für den Fall	entweder ... oder
obwohl	weshalb	gesetzt (den Fall)	oder
während	wobei	im Fall(e)	sondern
weil	wogegen	unterstellt	sowie
wenn	worauf	vorausgesetzt	sowohl ... als/wie (auch)
wie	zumal		und

Table 2: Examples of German conjunction types.

appear in. For a comprehensive discussion, see (Pasch et al., 2003). Figure 1 shows all the positions that an adverbial can potentially appear in. In the following sentences, the connective in the respective position is italicized.[2]

Null-stelle	Vorerst-position	Vor-feld	Nacherst-position	linke SK	Mittel-feld	rechte SK	Nach-feld
✓	✓	✓	✓	—	✓	—	✓

Figure 1: Possible positions for adverbial discourse connectives in German.

(13) Nullstelle.[3]
   Aber haben Sie schon mal ihre Disketten überprüft?
   'But did you check your floppy disks yet?'

(14) Vorerstposition[4].
   Sogar der Preis ist in Ordnung.
   'Even the price is ok.'

---

[2]Most examples are from (Pasch et al., 2003), if not marked otherwise.
[3]True pre-Vorfeld position. Can be filled in verb-first sentences, i.e., yes/no-questions and exclamations.
[4]Position for focus particles, immediately preceding the Vorfeld item.

4

(15) Vorfeld.
Allerdings habe ich das auch so gewollt.
'Admittedly, I wanted it this way.'

(16) Nacherstposition.[5]
Eine CD-Rom *dagegen* fasst bis zu 800 Megabyte.
'A cd-rom, on the other hand, can contain up to 800 megabytes.'

(17) Mittelfeld.
Das haben wir *aber* noch nie so gemacht.
'But we have never done it this way.'

(18) Nachfeld.
Das ist typisch deutsch *allerdings*.   (Callhome Training Corpus)
'Indeed, this is truly German.'

Nacherstposition.   German adverbial connectives can be distinguished into two basic groups, one whose members can appear in the post-Vorfeld position (Nacherstposition), and one whose members can't hold this position.

Of the adverbials that can not appear in Nacherstposition, by far the majority are pronominal adverbials, i.e., they contain a deictic element, usually a d-element: *davor* (before that), *dazu* (in addition), *stattdessen,* (instead) etc.

Pronominal and Relative Adverbs.   There are two special classes of adverbs in German that are also counted among the discourse connectives. They are the pronominal and relative adverbs (Pronominaladverbien, Relativadverbien). Pronominal adverbs consist of a deictic component and a relational component (usually a preposition), for example *da-vor, dem-nach, des-wegen, infolge-dessen*. Analogously, relative adverbs consist of a w-component and a relational component: *wo-vor, wes-wegen*.

The deictic or w-component fills one of the semantic argument slots since it is *referring*. It is, however, anaphorically linked to a clause or group of clauses that express a proposition. This is the reason why these adverbs, although really only semantically unary, are still considered discourse connectives. In the case of relative adverbs, the semantically coreferring clause is in addition syntactically required, since these adverbs belong to the class of postponers (a type of conjunctions). Furthermore, as documented in their name, these adverbs introduce relative clauses, which must as always be attached syntactically to something coreferring in a main clause. On the other hand, pronominal adverbs only have one syntactic argument, just like all other adverbial connectives. See section 3 for a more detailed discussion of anaphoric connectives.

2.3   Cross-linguistic Comparison of Conjunction Types

In English, where word order in subordinated clauses doesn't differ from the word order in main clauses, there is only a two-way distinction of conjunctions

---

[5]Following the Vorfeld, but preceding the finite verb. See below for details.

with regard to embedding. How can this view be reconciled with the 4 types of German conjunctions?

If we ignore word order in German, subordinating conjunctions proper and V2-embedders fall together into what are the subordinating conjunctions in English (because their internal argument is embedded into the external argument), whereas postponers and coordinating conjunctions (proper) merge into coordinating conjunctions, since there is no embedding in that case. This view has a good motivation because it captures the external syntactic properties of the conjunctions well (see section 4). It will make it possible to distinguish syntactically between subordinated clauses (containing a subordinating conjunction or V2-embedder), which can occupy a syntactic position within the main clause, and clauses introduced by coordinating conjunctions and postponers, which always have to follow the other clause and cannot be integrated into it.

However, it then remains to be explained in what way the connectives govern the word order in the embedded clause. We now see that word order cuts across syntactic classes, contra the traditional view (which would be that coordinators don't govern word order, but subordinators always demand verb-final word order[6]). In fact, there are two distinct problems here:

Question 1: Embedding and Embedded Word Order. How can it be syntactically explained that some embedders (the subordinating conjunctions) can only be followed by VF-order, and some (the V2-embedders) take only internal arguments with V2-order?

Question 2: Postponers. Coordinating conjunctions behave the same way as in English, in that they don't impose word order restrictions either way. That is, a coordinated clause exhibits whatever word order a non-coordinated clause in the same syntactic position would exhibit.

(19) <u>Geh</u> nach Hause *und* <u>mach</u> deine Hausaufgaben!   (V1)
      'Go home and do your homework!'

(20) Ich <u>koche</u> Mittagessen *und* er <u>bäckt</u> Kekse.   (V2)
      'I make lunch and he bakes cookies.'

(21) [Wir sehen uns nie], weil ich gerne lange <u>schlafe</u> *und* er gern frueh
      <u>aufsteht</u>.   (VF)
      '[We never see each other], because I like to sleep in and he likes to get
      up early.'

We will come back to both of these questions in section 4 on the syntactic analysis of connectives.

---

[6]In a transformational approach, this fact would be explained by the assumption that German V2-order is achieved by raising the verb to C. This raising is of course prevented if there is an overt complementizer (for example, a subordinator) in C, which leads to the underlying, verb-final word order.

3 Structural vs. Anaphoric Connectives

According to the PDTB, words (or phrases) are considered connectives in English iff they have two semantic arguments that are propositions, events or situations. There are, then, two types of connectives: the structural connectives, which have two local syntactic arguments that correspond to the two semantic arguments, and the anaphoric connectives, which take only one syntactic argument. The referent for the second semantic argument is derived anaphorically from the context. This has been shown extensively in Bonnie Webber's work (see Webber et al., 2003).

Trying to carry this argumentation over to the German case, one notices certain problems. Certainly, German is similar to English in the fact that connectives can take either one or two syntactic arguments. Conjunctions are structural, adverbials have only one local argument. However, not all German discourse connectives seem to be semantically binary. As introduced above (see section 2.2.2), two classes of connectives in German contain an internal part that expresses one of their arguments. These are the pronominal and relative adverbs.

As explained above, pronominal and relative adverbs consist of an anaphoric component (deictic or w-, respectively) and a relational component (usually a preposition). The anaphoric component fills one of the semantic arguments slots of the adverb word-internally. This has the effect that there is only one semantic argument left to be filled, which makes these connectives in principle unary.

However, it can be argued that the adverbs are still semantically binary relations. The relative adverbs have two syntactic arguments. Furthermore, their external syntactic argument is required to be coreferent with the w-component contained in the adverb. Thus, although the syntactic argument is not a direct semantic argument of the relation, it definitely specifies the exact referent for one participant in the relation, and can thus be regarded as an indirect argument.

The case of pronominal adverbs is less straightforward, because they take only one syntactic argument, which corresponds to the one open semantic slot. Although their second semantic argument slot is internally filled, it is filled with a deictic anaphor, similar in meaning to German *das* (this/that). The anaphor requires an antecedent, since although it is referential, only the antecedent will specify exactly what the anaphoric component is referring to.

Therefore, we can say about pronominal and relative adverbs that they are indeed two-place relations, for which one argument has to be determined anaphorically.

Having determined this, one immediately encounters another question: Is there any difference, then, between adverbial discourse connectives in general and pronominal and relative adverbs in German? It seems that the latter contain overt anaphoric elements, where the former only contain a covert anaphoric element. Naturally, the question for English is whether there are any connectives in which the anaphoric component is (synchronically or diachronically) overt. Obvious candidates are *therefore* (because of that), *thereby, thereafter*. These

Adverbials: S   VP   etc.

◇*conn* S*   ◇*conn* VP*

Subordinating Conjunctions: S   S

S↓ ◇*conn* S↓   ◇*conn* S↓ S↓

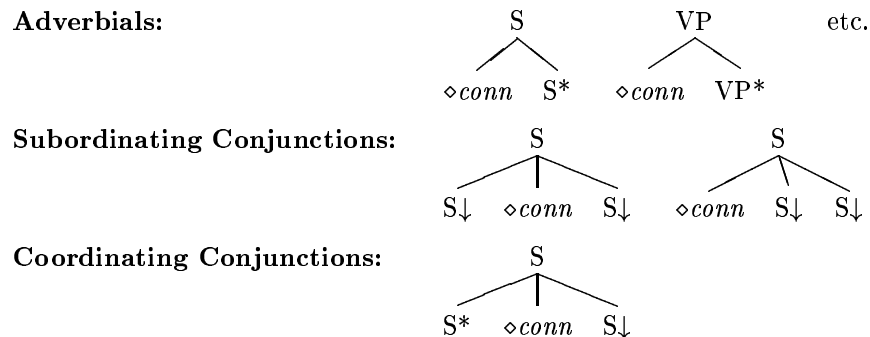**Coordinating Conjunctions:**   S

S* ◇*conn* S↓

Figure 2: Elementary TAG trees for English connectives.

can be considered overtly anaphoric only diachronically, though, since *there* does not mean *that* in modern English. Synchronically, I have not come across any overtly anaphoric English connectives yet.

To summarize, I have shown that the full classification of connectives considering their semantic and syntactic arguments includes 4 distinct types: connectives with 2 syntactic and 2 semantic arguments (most conjunctions), those with one syntactic argument, but 2 semantic arguments (most adverbials), those that have one syntactic argument and only one semantic argument, because the other semantic argument is already filled by an overt anaphor (like the German pronominal adverbs and *therefore*), and those connectives that have two syntactic arguments, but one of their semantic arguments is already filled by an anaphor, leaving only one semantic argument (like German relativ adverbs). It seems that connectives of this last type do not exist in English.

4   (TAG) Syntactic Analysis of Connectives

A TAG analysis for the three types of English discourse connectives has been proposed in the DL-TAG project (Forbes et al., 2001). Figure 2 shows the elementary trees associated with adverbials, subordinating conjunctions and coordinating conjunctions in English. I will discuss the corresponding German connectives in turn.

The analysis for adverbials is justified because they take only one syntactic argument, and the other semantic slot is filled anaphorically (see section 3). Additionally, all these discourse connectives are in fact *adverbials*, that is parts of speech that modify clauses. Auxiliary trees are used in TAG to formalize modification.

For German, this analysis can in principle be adopted as well. In German, just as in English, adverbial connectives have only one argument. Some ad-

justment might have to be made (i.e., some trees might have to be added) to
account for the greater number of position possibilities for German adverbials
within the clause they modify.

Similarly, the analysis for coordinating conjunctions can be adopted for German
coordinating conjunctions as well. In both languages, coordinating conjunctions
are not integrated into either conjunct, and can only appear between
the two conjuncts. Further, coordination can be seen as a modification (or
elaboration) of the first conjunct. Postponers are a special case, though. Their
internal argument must exhibit verb-final word order. In a TAG, this can be enforced
using features. What it makes clear, though, is that postponers still differ
from coordinating conjunctions substantially. An analysis more closely resembling
relative pronouns (considering that most postponers are *relative adverbs*)
may be granted. I will leave that for further research.

The analysis for English subordinating conjunctions can not be carried over
to the German case easily. Embedding in German is identified by the fact that
the connective plus its internal argument together can fill a syntactic position
in the external argument (for example, the Vorfeld position (9) or a Mittelfeld
position). This can not be captured by the trees for English subordinating
conjunctions given above, since the external clause is not necessarily a complete
clause – rather, the connective plus its internal argument is clearly a constituent
in the external argument.

Clearly, we are experiencing a mismatch of syntactic and semantic structure
here. While semantically, the two propositions expressed by the two clauses are
the arguments of the connective, the structure corresponding to the connective
plus its internal argument are syntactically a constituent of the external clause.
Certainly, the subordinated clause is not a syntactic *argument* of the external
clause. Rather, because the subordinated clause behaves similarly to adverbial
modification, the connective must be analysed as a modifier of its external
argument. In TAG terms, this corresponds to an auxiliary tree. Thus, we propose
the trees in figure 3 for the syntactic analysis of German subordinating
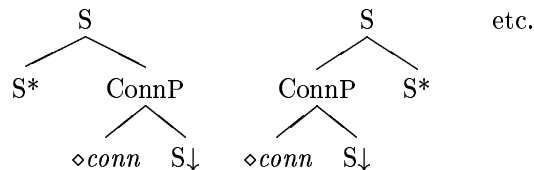conjunctions.

Figure 3: Elementary TAG trees for German subordinating conjunctions.

The given trees are valid for both subordinating conjunctions and V2-embedders,
the difference being in the word order requirements that are posed on the
S↓ argument. These trees are similar to the analysis for prepositions. Prepositional
phrases, as well, are syntactically embedded into the noun phrases they
modify. Another syntactic analogue are relative clauses, which are in many ways
similar to embedded clauses introduced by connectives. The difference is that

here, a clause (verbal projection) is modified, whereas relative clauses modify nouns.

5 Attributions

When annotating discourse connectives and their arguments (as in the PDTB), annotators often disagree on the scope of the arguments in case there is a verb of *attribution* in the sentence. For this reason, in the English annotation, it has been found useful to annotate attributions separately. Each argument of a discourse connective, as well as the discourse connective itself, is usually attributed to at least one participant in the discourse context. Such a participant can be the author of the text, but also people or institutions that are directly or indirectly quoted in the text.

For German, a similar effect can be found. In general, arguments as well as the relation expressed by the connective itself can be attributed to a participant. I have found examples where both arguments and the discourse relation are all attributed to the author of the text (22), and other examples where the relation as well as both arguments are attributed to some speaker that is mentioned in the text (23, note the subjunctive in the arguments that gives away the scope of the attribution).

(22) Weil es den internationalen Normen nicht mehr entspricht, plant die Landesregierung ein neues.   (Frankfurter Rundschau, Feb. 2, 1998) 'Because it doesn't conform to the international norms any more, the state government is planning a new one.'

(23) [*In Teheran hieß es*, die Frau sei zum Auspeitschen verurteilt worden;] die Todesstrafe sei ihr erspart geblieben, weil sie nicht verheiratet sei.
        (Frankfurter Rundschau, Feb. 2, 1998)
'[*In Teheran they said* that the woman was sentenced to whipping;] she was saved from the death penalty, because she isn't married.'

However, so far I have not come across any mixed versions, that is, any examples where one or both arguments are attributed to one participant, and the relation to another participant. A convincing example for English is quoted in (Dinesh et al., 2005, their number (8)):

(24) *The current distribution arrangement ends in March 1990,* <u>although</u> Delmed said **it will continue to provide supplies of the peritoneal dialysis products to National Medical**, the spokeswoman said.

Here, the internal argument (in bold) is attributed to Delmed, and the external argument (in italics) as well as the relation itself to the spokeswoman.

It is an open question whether such examples exist in German, and if so, if they have any special semantic properties.

6 Adverbials' External Arguments

In English, some discourse adverbials pose special requirements on their non-local argument. For example, *instead* requires its external argument to be negated in some way. This can be accomplished by explicit negation, by the inclusion of a negative quantifier (such as *no*, *few*), or modal, or even by the use of a verb or other constituent that has a negative implication (see Webber et al., 2003).

(25) John **refused** to eat the vegetables. *Instead*, he only wanted icecream.

(26) DSM's position with regard to a worldwide paper trail, however, is that it is prone to fraud and **impossible** to police. *Instead*, DSM favors a labeling system that is meaningful to consumers [. . .].

The German connective parallel to English *instead* is *stattdessen*. *Stattdessen* can be licensed under explicit as well as implicit negation of the external argument.

(27) Peter hat sich **geweigert**, das Gemüse zu essen. *Stattdessen* will er nur Eis.
'Peter **refused** to eat the vegetables. *Instead*, he only wants icecream.'

(28) Kleine Kinder sind **ungern** den ganzen Tag zuhause.
Little children are **unwillingly** the whole day at home.
Stattdessen rennen sie lieber draußen herum.
Instead, run they rather outside around.
'Little children don't like to be at home all day. Instead, they prefer to run around outside.'

The coordinating conjunction *sondern* corresponds in meaning to *instead*, as well. In contrast to the English adverbial, *sondern* usually requires *explicit* negation, i.e. negation by *nicht* (not), adverbs like *nie* (never), *kaum* (rarely) etc., or by the negative quantifier *kein* (29). A verb or other word with a "negative implication" isn't sufficient to license *sondern* (30).

(29) Er ist **kein** Junggeselle, *sondern* verheiratet. 'He is not a bachelor, but married.'

(30) * Er ist **unverheiratet**, *sondern* Junggeselle.
int.: 'He is not married, but a bachelor.'

However, in naturally occurring speech, such examples can sometimes be found:

(31) Es [. . .] **stand** immer **außer Frage**, daß wir heiraten, *sondern* es war nur die Frage so wann. (HUB Corpus)
'It was always **out of the question** that we will get married, *(but)* it was only the question, when.'

It is open for further research to determine which other connectives pose special requirements on their external argument. Furthermore, it has to be determined by what process connectives can impose their restrictions on the syntactic form of their external argument, especially in the case of adverbials when the external argument is only related anaphorically (not structurally) to the connective itself.

References

Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. 2005. Attribution and the (Non-)Alignment of Syntactic and Discourse Arguments of Connectives. In *Proceedings of the ACL Workshop on Frontiers in Corpus Annotation II*, Ann Arbor, Michigan.

Katherine Forbes, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind Joshi, and Bonnie Webber. 2001. D-LTAG system − Discourse parsing with a Lexicalized Tree Adjoining Grammar. In *Proceedings of ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*, Helsinki, Finland.

Renate Pasch, Ursula Brauße, Eva Breindl, Ulrich Hermann Waßner. 2003. *Handbuch der deutschen Konnektoren.* Berlin: Walter de Gruyter.

Rashmi Prasad, Eleni Miltsakaki, Aravind Joshi, and Bonnie Webber. 2004. Annotation and Data Mining of the Penn Discourse Treebank. In *Proceedings of ACL Workshop on Discourse Annotation*, Barcelona, Spain.

Bonnie Webber, Matthew Stone, Aravind Joshi, and Alistair Knott. 2003. Anaphora and discourse structure. *Computational Linguistics*, 29(4):545–587.

Corpora

Krisjanis Karins, Robert MacIntyre, Monika Brandmair, Susanne Lauscher, and Cynthia McLemore. 1997 *CALLHOME German Transcripts.* LDC97T15. FTP file. Philadelphia: Linguistic Data Consortium.

LDC. 2003. *1997 HUB5 German Transcripts.* LDC2003T03. FTP file. Philadelphia: Linguistic Data Consortium.