Discourse particles and routine formulas in spoken language translation

Manfred Stede and Birte Schmitz

Technische Universität Berlin

Sekr. FR 5–12

Franklinstr. 28/29

10587 Berlin, Germany

Email: {stede|birte}@cs.tu-berlin.de

Abstract

The problem of polysemy has received much attention when dealing with content-words, but it is just as difficult for *discourse particles*: In spoken language, they often perform various functions for dialogue management, rather than contributing to propositional content. Different languages have evolved different conventions on using such particles, which renders the task for spoken language translation quite difficult. We focus here on particles in German, suggest a framework for representing their roles in utterances, and sketch an approach for adequately translating them into English.

1 Overview

Discourse particles at first sight seem to be innocent little words, but they can pose significant problems for automatically processing spoken language. Their abundance (*types* and *tokens* alike) varies notably from language to language. In this paper we are concerned with German and English; in the former, a particularly large number of such particles is in use. In the next section, we demonstrate that these particles can be quite problematic in translation, drawing on extensive corpus analyses we performed in the VERBMOBIL project [Wahlster 1993]; thus our examples are all from the domain of appointment scheduling. Thereafter, we discuss the consequences for automatic spoken language translation (SLT) and suggest a framework of *discourse functions* to adequately represent the role of the particles within utterances. We then point out that certain *routine formulas* can be tackled with the description of discourse functions as well. Finally, we discuss how our inventory of discourse functions can be used to improve translation quality in an SLT system such as VERBMOBIL.

2 Discourse particles: a rich source of ambiguity

Discourse particles are words that are not uttered because of their propositional content, but because of some pragmatic function for the ongoing discourse. Imagine, for instance, utterance (1) in the midst of a discussion; the *right* at the beginning serves mainly to signal turn-taking and initiating some kind of break in the conversation.

(1) *Right, now let's turn to the next topic.*

The exact function of a discourse particle is often difficult to determine, though, and thus the need for disambiguation arises: In the German utterance (2), which can be roughly translated as *So we're finished*, the *ja* can be a mere filler that smoothens the intonation contour, or it can mark the overall information as *given*, e.g. in a situation where the participants have just collectively closed off the final topic.

(2) *Dann sind wir ja fertig.*

The disambiguation problem is aggravated considerably by the fact that the vast majority of words that can be used as discourse particles also have a "standard", lexically-determined, reading.[1] In (1), the *now* can be read as a temporal adverb or as a semantically empty particle. Similarly, in (2) the *dann* is most likely just an 'uptake' particle, but in general it is a temporal adverb as well.

We will follow the terminology of Hirschberg and Litman [1993], who distinguish between the *sentential* usage and the *discourse* usage of such words. In these terms, the problem is that many particles have a sentential usage, but it may very well be irrelevant in certain contexts, where only the discourse usage of the particle is relevant — and often that usage has no systematic relationship to the sentential usage.

---

[1] Exceptions are, for instance, the English *oh* or the German *ach*, which are always discourse particles.

Several studies of English discourse particles have suggested that utterance-initial position is the central criterion to distinguish sentential from discourse reading (e.g., Reichman [1985]), and certainly, adopting such a criterion makes it much easier to gather and evaluate the data (cf. Byron and Heeman [1997]). The position-criterion, however, holds for a certain class of discourse particles at best. Hirschberg and Litman [1993] give some counter-examples to the hypothesis. Furthermore, it does not travel well to the German language, despite the fact that Ripplinger and Alexandersson [1996] also give the utterance-initial criterion for identifying German discourse particles, such as *denn, also, ja*, in the VERBMOBIL corpus. In fact, these and other particles can occur almost everywhere in the utterance.

For this and other reasons, the class of discourse particles is a heterogeneous one and very difficult to demarcate in syntactic or other formal ways. Schiffrin [1987, p.31ff] develops a lengthy definition of *units of talk* and then treats discourse particles (*markers* in her terminology) as bracketing such units. For our purposes here, we do not attempt to give formal criteria for what constitutes a discourse particle; instead we are content with the functional, dual-usage description. This results from taking a rather broad perspective on the range of functions of discourse particles, which will be developed in section 3.

For the task of translating utterances containing discourse particles, the monolingual ambiguity problem is extended by that of finding an appropriate translation — and different languages have developed quite different conventions for using particles. German is known to offer an especially wide range of discourse particles, and therefore it is not surprising that many of them do not have any straightforward English translation at all — instead, their function in discourse needs to be signalled by different means. And in many cases a particle is best dropped from the translation process altogether, if its existence is due merely to certain conventions in spoken German, which do not carry over to the target language.[2] The problem is, given a particular utterance containing a particle, to tell which case is at hand.

---

[2]Furthermore, the problem is amplified by the fact that German particles can be combined to form conglomerates, as in *Wir sollten wohl doch noch mal einen Termin ausmachen*. To what extent these can be analyzed and translated compositionally is an open question. In the present paper, though, we deal only with individual particles.

2.1 German examples

The VERBMOBIL system currently operates with a German vocabulary of 2300 words, and among these we have identified 49 particles that cause problems of the kind just described. (This figure does not include modal and other particles that can pose significant problems for translation but do not have a discourse usage.) Although we have not performed large-scale frequency tests, some initial counts suggest that a typical dialogue from the VERBMOBIL corpus, which consists of about 15 turns, contains 20 to 30 occurrences of such particles. For this paper, we have chosen 7 particles to illustrate the situation for disambiguation and translation.

doch If *Doch!* is the sole response to an utterance, $doch_1$ denies what the other participant has just said and at the same time reaffirms the (opposite) opinion of the speaker. In English, one would say something like *Yes it is!* (or use the more specific verb in question). However, when *doch* is used within a sentence, it has at least the following three functions. In *Das ist doch klar*, $doch_2$ signals that the proposition is assumed to be shared knowledge, or self-evident. A suitable English rendering is *That's clear anyway*.

(3) *Lassen Sie uns doch einen Termin ausmachen.*

But when a sentence like (3) introduces an exchange, $doch_3$ merely smoothens the utterance and should not be translated at all, since English does not offer a corresponding word. Thus, a translation is *Let us arrange an appointment*.

(4) *Dann nehmen wir DOCH Dienstag.*

Finally, in an utterance like (4), where $doch_4$ is prosodically marked, it signals a return to a previous state of the negotiation: Tuesday had been suggested earlier but was rejected, and now the rejection is taken back. Again, there is no equivalent English word; instead, a speaker can signal the reversal of her position by saying, for instance, *All right, so we DO take Tuesday*.

noch Another example is *noch*, which can be a semantically empty smoothening particle, or a focus particle meaning *another*.

(5) *Wir müssen noch einen Termin ausmachen.*

Sentence (5) is ambiguous between *We have to arrange an appointment* and *We have to arrange another appointment*. Often, prosody indicates the distinction [Bos et al. 1995]; otherwise, analysis of the preceding context is necessary: Has a different appointment been scheduled already, so that another one can be dealt with now?

also The "literal" meaning (according to dictio-

naries) of *also₁* is *therefore* or *so*.[3] In a related usage, *also₂* can introduce a reformulation or specification of information:

(6) *Treffen wir uns am Dienstag, also am dritten Juni.*

Sentence (6) can be translated as *Let us meet on Tuesday, hence on the third of June* (the more formal *therefore* would be misplaced here).

(7) *Also, da muss ich mal schauen.*

However, in (7), *also₃* does not contribute to sentence meaning but merely signals that the speaker is taking her turn. In English, this is typically performed by *now, ...* or *well, ...,* etc.

bitte Sometimes, the dialogue act of the preceding utterance can be the decisive piece of information for disambiguation. The single-word utterance *bitte* is ambiguous between *you're welcome!* and *excuse me?*, asking the other participant to repeat the utterance. Now, in case the dialogue act of the preceding utterance is THANK (which abstracts over the various linguistic means of expressing gratitude), *bitte* is translated as *you're welcome.* (Prosodic information can also help to distinguish the two forms of *bitte*, but is not always reliable.)

beziehungsweise This is a curious German word that can be employed for quite a variety of purposes.

(8) *Mir passt es am Dienstag beziehungsweise am Freitag gut.*

For example, in (8) the speaker lists two alternatives, so that *beziehungsweise₁* is best translated as *or*, thus: *For me, Tuesday or Friday is fine.*

(9) *Ich bin am Dienstag und Mittwoch, also am dritten beziehungsweise am vierten, in Hamburg.*

In an enumeration with specification, such as (9), the English *and ... respectively* is an adequate translation for *beziehungsweise₂*; only in this case, a relation between different elements is verbalized, which is part of the "literal" meaning of *beziehungsweise*. It requires a certain amount of reasoning, though, to notice this relation between 'Tuesday' and 'third', and between 'Wednesday' and 'fourth', so that the translation *I am in Hamburg on Tuesday and Wednesday, hence on the third and fourth, respectively* is achieved.

(10) *Am Montag bin ich in Frankfurt, beziehungsweise in Eschborn.*

Often, however, *beziehungsweise₃* merely starts a self-correction, specification, or reformulation, best translated as *or rather*, as in (10): *On Monday I am in Frankfurt, or rather in Eschborn.* (Eschborn is a smaller town neighbouring Frankfurt).

---

[3]Notice that German *also* has no relationship to English *also* — they are genuine *false friends.*

vielleicht Dictionaries give the meaning *maybe* for this adverb, but in spoken language it is often produced merely as a filler word, partly indicating vagueness, partly buying time for the speaker to do more thinking. In (11), the *vielleicht₂* is to be treated as such a filler; the literal translation *Then let us maybe say Tuesday* is unnatural.

(11) *Dann sagen wir vielleicht am Dienstag.*

ja As a complete utterance, the meaning of *ja* is quite significant ('yes'), but as a little word accompanying utterances it can be used for many, often less significant, purposes. For instance, it can serve as a simple uptake, similar to *also₃*; or it can be a plain filler that should not be translated at all; or it can indicate that the speaker assumes something to be presupposed by the hearer as well — in a negotiation, someone would say (12) if he had already given the information *On Monday I am in Hamburg* earlier.

(12) *Montag bin ich ja in Hamburg.*

2.2 English examples

In English, particles are generally less frequent than in German. Correspondingly, using particles for discourse purposes is also less common than in German. While we have not done comprehensive corpus studies on English data yet, the limited role of particles can be inferred from the fact that the research literature often deals merely with those occurring utterance-initial (e.g., [Byron, Heeman 1997]). Especially the *structuring* particles are well-studied; for instance, *by the way* and *but* can signal the beginning of a digression, and *anyway* the return to a previous topic.

Other prominent utterance-initial particles are those that signal turn-taking (we call them 'uptakes'), as in (13) and (14), taken from the English VERBMOBIL data. In (14), the *well* is not just a neutral uptake but also indicates reservation.

(13) *Alright, why don't you come to my office.*

(14) *Well, the morning of the eighteenth is bad.*

Some studies look at particles in other positions as well. For instance, Redeker [1990] considers utterance-final tags like *okay?* or *right?*, which serve to elicit acknowledgement from the listener, and *comment-clauses*, which can be used as 'enquoting devices' to signal the beginning of quoted speech.

Other particles, which we also find in the English VERBMOBIL data, include *repair markers* such as the *oh* in (15) and the *no* in (16). Incidentally, (16) is an example for a correction that does not pertain to the part of the utterance immediately preceding the correction marker.

(15) *I'm going to be in Cleveland on the sixth,*

5

```
                              discourse function
        _____|_____
       |                    |                 |                  |
   structuring      coherence-marking      attitudinal       smoothening
    __|_|_|__                            __|_|_|__               |
   |  | | |  |                          |  | | |  |             |
 push pop uptake check repair   positive negative indifferent given surprise   filler
```
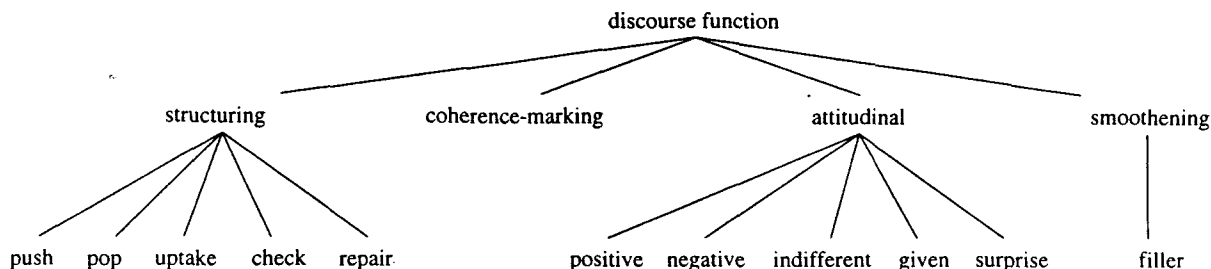
Figure 1: Taxonomy of discourse functions

how about sometime, oh, actually maybe this week isn't going to be so good.

(16) *I'm on vacation from the second to the twelfth, I'm going to see the Grand Canyon, no, actually to the fifteenth.*

2.3 Consequences for automatic translation

Summarizing the examples given above, a particle can have

- a "literal" lexical translation, where semantics (truth conditions) is relevant (there can still be ambiguity between several sentential readings),

- a "non-literal" lexical translation, where pragmatic intent is relevant (again, there can be ambiguity between several discourse readings),

- a non-lexical translation, i.e., it is rendered by a syntactic or intonation feature,

- a zero translation.

For some particles, only a subset of these cases is possible, but others can, depending on context, fall into any of these groups, such as the notorious *ja*.

Moreover, we have seen that a variety of knowledge sources is required to find the adequate translation:

- syntactic environment (e.g., for *ja*),

- semantic types,

- previous utterance (e.g., for *doch*),

- preceding discourse (e.g., for *noch*),

- world knowledge and/or inferences (e.g., for *beziehungsweise*),

- prosody (e.g., for *bitte*),

- genre-specific conventions (see example below in section 4).

Given this situation, it is unrealistic to expect that a set of standard transfer rules is sufficient to provide adequate translations of particles on a lexical basis. Instead, a "deeper" analysis is required, which extracts sufficient contextual information and also, if necessary, can bring world knowledge into play. Similarly, on the side of the target language, the best realization corresponding to a source language particle is not always straightforward to determine. This is another argument against relying solely on transfer rules, especially if the realization is an intonation feature or syntactic tag whose position is not immediately related to that of the source language particle.

3 Discourse functions

In VERBMOBIL, the "deep" analysis is undertaken in the Context Evaluation (ConEval) module, which constructs a conceptual representation (based on a domain model coded in a description logic language) from the output of the syntactic/semantic analysis module. The conceptual representation is used to compute the dialogue act [Schmitz, Quantz 1995] and to perform various disambiguations [Stede et al. 1996], which can involve inferences. For example, we reason with date expressions to determine whether one date is a specification of another, or a separate one (which is sometimes important for disambiguation).

In our conceptual representation, the discourse particles (in their "pragmatic reading") are represented by labels signifying their *discourse function*. Figure 1 shows our current taxonomy of discourse functions (an extension of the one given in [Schmitz, Fischer 1995]), which has resulted from extensive corpus analyses of the VERBMOBIL data and from examining the relevant research literature. In the following, we briefly explain the various functions.

Structuring These functions have received the most attention in the research literature. PUSH and POP mark the beginning of a sub-topic or di-

gression, and the return to the previous topic, respectively. (Examples: *by the way–anyway / davon abgesehen–wie auch immer*). With an UPTAKE, the speaker signals a turn-taking at the beginning of a turn and a turn-holding within a turn. It also can help the hearer adapting to the acoustic properties of the speaker's utterance without losing information. (Examples: *all right, now / ja, also*). CHECK is a turn yielding signal, prompting the dialogue partner to respond. By using a CHECK, the speaker often seeks approving feedback from the hearer. (Example: *isn't it? / oder?*). REPAIR indicates problems in planning and performing the output, signals a new start, and thereby is also a turn-holding signal. (Examples: *I mean, sorry / ach nein, äh*).

Coherence Marking Some particles can be employed to facilitate the embedding of the utterance within the context, and to check the common basis of the participants. (Example: *doch, schon*).

Attitudinal We borrowed this label from Allwood et al. [1992]. While English often uses verbs for these purposes, German also offers a range of particles for speakers to convey a POSITIVE (example: *gern*), NEGATIVE (example: *leider*), or INDIFFERENT (example: *ruhig*) attitude towards the propositional content in their utterance, or towards the last utterance of the dialogue partner. In addition, the propositional content of the speaker's own utterance can be marked as presupposed, or GIVEN (examples: *ja, doch*). And, some particles indicate SURPRISE at an utterance made by the partner (example: *oh*).

Smoothening Also especially in German, particles often help to create an overall appropriate intonation contour, and at the same time can serve to express cooperativity and politeness (Examples: *denn, doch*). Specifically, FILLERS allow the speaker to plan the output, avoid undue pauses, and help to hold the turn. (Examples: *ich würde sagen, äh*).

4 Routine formulas

We pointed out that the particles investigated here have at least one reading in which the *discourse usage* is central to their usage, and not semantic contribution to propositional content. This difference points to the notion of "idiomatic" meaning, and — not surprisingly — the discourse functions introduced above can often also be realized by idiomatic phrases. Without going into detail here, we merely give a few examples, again taken from the VERBMOBIL domain. In all these cases (and many others), the "literal" compositional meaning is not the point of using the phrase, and they typically cannot be translated word-by-word.

As **fillers**, we often find phrases like *Ich würde denken, ...* or *Ich muß sagen, ...* In English, the translation *I must say, ...* is not wrong but not conventionally used in this context. Similarly, the German *Wenn ich da mal nachsehe, ...* should not be translated preserving the conditionality, hence *If I look this up, ...* but by the common phrase *Let me see, ...*

The function **check** can be realized by phrases like *Sehe ich das richtig?* which also should not be translated literally (*Do I see that correctly?*) but by a conventional phrase such as *Am I right?*

repair markers can also be phrasal, as in *X, oder besser gesagt, Y* or in *X, nein, ich wollte sagen Y*. Again, literal translations should give way to conventionalized English formulas, hence *X, no, I wanted to say Y* is less felicitous than *X, no, I meant Y*.

5 Towards automatic translation

Since the problems associated with discourse particles are largely absent when processing written language, computational linguistics has for most of its history not dealt with these problems. In SLT, however, they cannot be avoided, especially when working with a language rich in particles, such as German. Given the youth of the field, plus the fact that particles at first sight do not exactly seem to be the most important challenge for translating spoken language, it comes as no surprise that there are no satisfactory solutions in implemented systems yet.

In the VERBMOBIL prototype that was completed last year, a number of particles are considered ambiguous between scopal/modal/focusing adverb on the one hand, and "pragmatic adverb" on the other. This class of "pragmatic adverbs" loosely corresponds to the "discourse usage" we have investigated above. The translation framework of VERBMOBIL is strongly lexeme–based; thus, for any particle in the German source-utterance, the transfer component seeks a corresponding English word on the basis of the reading determined. Typically, the ConEval module is asked to determine the class of a particle, wherupon transfer chooses a target word. As an exception, in some contexts a pragmatic adverb is suppressed in the translation.

This procedure is a start, but it cannot deal with all the facets of meaning found in discourse particles, as outlined above. On the basis of corpus studies, both [Schmitz, Fischer 1995] and [Ripplinger, Alexandersson 1996] already demonstrated that many German particles have a whole range of English correspondents, of which VERBMOBIL at present manages only very few.

To improve the translations, for the second phase

of the VERBMOBIL project we propose to build upon the framework of discourse functions. The purpose of computing discourse functions in *analysis* is twofold: it supports disambiguation (not only of the discourse particles, but also of the surrounding words) and computation of the dialogue act underlying the utterance; and it helps in segmentation, i.e., breaking an utterance into portions that serve as complete units for further processing. In *translation*, the information on discourse function is important for deciding whether to translate a particle at all, and how to do that: by inserting a corresponding target language particle, or by modifying the syntactic structure or intonation contour of the target utterance.

Given the wide variety of information required for determining discourse functions (listed in section 2.3), the task is best performed in tandem with building up the conceptual representation of the utterance, i.e., in the ConEval module. The decision as to what discourse function to associate with a particle is seldom a strict one (not even for the human analyst). Instead, the different clues from syntax, semantics, prosody, and world knowledge are typically weak and have to be weighted against each other in the light of the complete utterance. Therefore, we tackle the problem with the same mechanism we use for identifying dialogue acts: a set of weighted default rules, implemented in FLEX [Quantz et al. 1996] as an extension to the standard description logic language. The rules are matched against the utterance representation, and the accumulated weights decide on the most likely discourse function. We are currently in the process of defining this rule set.

The result will be more fine-grained information on discourse particles than is available now in the system. The transfer and generation modules can use the discourse function to decide whether a lexical correspondent should be produced in the target language, and if so, which one, and at what position of the utterance. Particles that are mere fillers can be removed entirely from the translation, and similarly those particles that are used to smooth the intonation contour in German. Whether restarts and self-repairs get translated or are merged into a single coherent utterance, is an open question. In many cases, it would not be difficult to replace the "corrected" portion of an utterance with the portion that "overwrites" it, thereby sparing the hearer from reworking the correction herself.

As for routine formulas, they first of all cause the standard problems of idiomatic phrases: they need to be recognized as a single unit of meaning, so that they can be translated *en bloc*. This presupposes lexical representations that adequately describe the possible variants of the expression, e.g., whether additional modifiers may be inserted into a phrase, etc. When processing written language, this is difficult enough — with speech and the additional uncertainties of word recognition, the problems are even harder. For the time being, a comprehensive treatment of routine formulas and other idioms does not seem feasible.

Regarding the overall system architecture, the deep-analysis phase, as we have described it, need not be necessary for each and every utterance — if the input allows for a standard transfer-based translation (e.g., because it doesn't contain ambiguous particles), that will typically be sufficient. This essentially amounts to a mixed-depth analysis in the translation process — an important question that we cannot discuss further here.

6 Summary

Discourse particles and routine formulas in spoken utterances cannot be translated on a simple lexeme-to-lexeme basis. We have proposed a taxonomy of *discourse functions* to represent the pragmatic impact of such particles and formulas. There is, still, no 1:1 mapping between particles/formulas and discourse functions in analysis, nor between discourse functions and their realization in the target language. Therefore, we use a "deep" utterance representation of dialogue act and propositional content, into which discourse functions are integrated. In analysis, the deep representation holds all the information required for successful processing; the transfer and generation components can then decide whether discourse functions get realized in the target language, and if so, by what means. This decision can be made in the context of the target-language utterance.

References

W. Abraham (ed.). *Discourse particles*. Amsterdam: John Benjamins, 1991.

J. Allwood, J. Nivre, E. Ahlsen. "On the semantics and pragmatics of linguistic feedback." In: *Journal of Semantics* 9:1–26, 1992.

J. Bos, A. Batliner, R. Kompe. "On the use of Prosody for Semantic Disambiguation in VERBMOBIL". Verbmobil Memo 82, DFKI Saarbrücken, 1995.

D. Byron, P. Heeman. "Discourse marker use in task-oriented spoken dialog." In: *Proceedings of the 5th*

European Conference on Speech Communication and Technology (Eurospeech), Rhodes/Greece, 1997.

J. Hirschberg, D. Litman. "Empirical studies on the disambiguation of cue phrases." In: *Computational Linguistics* 19 (3), 501–530, 1993.

J.J. Quantz, G. Dunker, F. Bergmann, I. Kellner. *The FLEX System.* KIT Report 124, FB Informatik, TU Berlin, 1996.

R. Reichman. *Getting computers to talk like you and me.* MIT Press/Bradford Books, 1985.

B. Ripplinger, J. Alexandersson. "Disambiguation and translation of German particles in Verbmobil." Verbmobil Memo 70, IAI/DFKI Saarbrücken, 1996.

D. Schiffrin. *Discourse markers.* Cambridge University Press, 1987.

B. Schmitz, K. Fischer. "Pragmatisches Beschreibungsinventar für Diskurspartikeln und Routineformeln anhand der Demonstratorwortliste." Verbmobil Memo 75, TU Berlin/Universität Bielefeld, 1995.

B. Schmitz and J.J. Quantz. "Dialogue-Act Type Recognition in Automatic Dialogue Interpreting." In: *Proceedings of the The Sixth International Conference on Theoretical and Methodological Issues in Machine Translation (TMI)*, Leuven, 1995.

M. Stede, U. Küssner, J. Quantz, B. Schmitz. "Disambiguation as part of semantic evaluation in Verbmobil." In: C. Boitet (ed.): *Proceedings of the Workshop on Multimodal Interactive Disambiguation (MIDDIM-96)*, Grenoble, 1996.

W. Wahlster. "Verbmobil: Translation of face-to-face dialogues." In: *Proceedings of the Third European Conference on Speech Communication and Technology*, Berlin, 1993.