Speech acts that support other speech acts

Katja Jasinskaja KATJA.JASINSKAJA@UNI-KOELN.DE
University of Cologne

Frank Zickenheiner FRANK.ZICKENHEINER@UNI-BONN.DE
University of Bonn

Editor: Name Surname

Submitted 06/2024; Accepted MM/YYYY; Published online MM/YYYY

Abstract

Theories of discourse structure and existing discourse structure annotation schemes (e.g. Mann and Thompson, 1988; Sanders, 1997; Asher and Lascarides, 2003; Webber et al., 2019) often make a distinction between *propositional*, *subject-matter*, or *semantic* coherence relations on the one hand, and *speech-act-level*, *presentational*, or *pragmatic* relations, on the other. While there have been several convincing attempts to circumscribe the space of all possible propositional relations and to subdivide it into theoretically motivated subcategories (Sanders et al., 1992; Kehler, 2002), to date there is no comparable comprehensive taxonomy for speech-act-level relations. This paper develops a fragment of a such taxonomy, which describes what we call *support relations*—relations that connect two speech acts iff one of them fails to achieve its goal and the other helps achieve that same goal, as for instance, in *Evidence* relations, where one speech act makes the proposition asserted in the other more believable. We provide conceptual motivation for the proposed categories grounded in isights from sociolinguistic, psychological, and philosophical studies of human communication and illustrate the categories with examples from naturally occurring discourse, some of which do not fit easily into any existing classifications.

Keywords: coherence relations, speech acts, communicative goals, dialogue, monologue, belief

Acknowledgments

The research reported in this paper has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 281511265 – SFB 1252 "Prominence in Language" in the project C06 at the University of Cologne, Department of German Language und Literature I. We would also like to thank the audiences of the workschop on *At-Issueness, Scope, and Coherence (ASC)* 2019 in Cologne, the SPAGAD 2 workshop on *Speech Acts in Grammar and Discourse* 2020 at ZAS, Berlin, and the 24th *Szklarska Poręba Workshop on the Roots of Pragmasemantics*, especially Matej Drobňák, for inspiring discussions of previous versions of this paper.

1. Introduction

The central assumption of a large group of approaches to discourse structure (Mann and Thompson, 1988; Sanders et al., 1992; Kehler, 2002; Asher and Lascarides, 2003) is that a discourse is coherent to the extent that hearers or readers are able to connect the sentences and larger discourse units it consists of with meaningful links—coherence relations (*alias* rhetorical relations or discourse relations, see Jasinskaja and Karagjosova, 2020, for an overview). While details might differ between the approaches, many of them make a distinction between *propositional*, *subject-matter*, *event-level*, or *semantic* relations on the one hand (1-a), and *speech-act-level*, *presentational*, or *pragmatic* relations, on the other (1-b). So in (1-a), a causal coherence relation holds between the described events of pushing and falling, whereas in (1-b) the causal relation holds between the speech act of asking the question and the fact that there is a good movie on (Sweetser, 1990).

(1) a. Max fell. John pushed him.
b. What are you doing tonight? Because there's a good movie on.

The focus of this paper is on relations of the second kind—those that connect speech acts in discourse. While a number of speech-act-level relation types have been identified in previous studies and existing discourse structure annotation schemes (Asher and Lascarides, 2003; Sanders et al., 2009; Webber et al., 2019), to date there is no comprehensive taxonomy of such relations. The question that the research reported in this paper is trying to answer is what kinds of coherence relations between speech acts exist, *should* exist, and are *theoretically possible*. The emphasis is here on 'should' and on 'theoretically possible', as our approach is not empirical or exploratory in the sense of looking for speech-act-level relations in naturally occurring discourse and trying to define fitting categories for the cases we find. Our goal is to find out how speech acts should relate to each other based on a theoretical understanding of what a speech act is, what definitional properties it has, and how those properties impose constraints on speech act combinatorics.

In this paper we argue that coherence relations between speech acts are best understood in terms of relations between their communicative goals. At this point, we only develop a fragment of the taxonomy, focusing on what we call *support relations*, i.e. relations where one speech act is produced as a means to successfully achieve the communicative goal of another. The class of support relations overlaps several prominent relation classes familiar from previous studies, in particular, presentational relations in Rhetorical Structure Theory (RST, Mann and Thompson, 1988) and some kinds of self-repairs (Levelt, 1983; Clark, 1994). However, it also includes cases that do not seem to fit comfortably into any existing classifications or annotations schemes. The relation between (2-b) and (2-c) is a case in point:

(2) a. "Pa told Peter he wanted *him* to be chairman."
b. "Sure he did.
c. If you don't believe me, ask Peter."[1]

1. From *Night over Water* by Ken Follett.

The speech act in (2-c) is a directive, which asks the addressee to seek evidence for the proposition asserted in (2-b). Its purpose is to make (2-b) more believable, but it does so without directly providing evidence, and therefore does not fall under the standard definitions of *Evidence* relations. The focus of relational theories of discourse structure has traditionally been on written texts that consist predominantly of assertions, and even though extensions of the approach to dialogue have been proposed (Taboada, 2004; Asher and Lascarides, 2003; Lascarides and Asher, 2009), the coverage of non-assertive speech acts remains limited. The framework developed in this paper naturally accommodates speech acts of all types.

In section 2, we argue that coherence between speech acts is governed by a different set of principles and should not be treated by analogy with propositional-level or semantic relations. We discuss previous work on speech-act-level coherence in dialogue and show how those findings can be applied to monologue. Finally, the notion of support relation is introduced in this section. In section 3, we define goals of speech acts and identify necessary subgoals that a speech act must achieve in order to be successful. Section 4 presents a fine-grained classification of support relations along two main dimensions: (a) the subgoal at which the initial speech act fails, which the supporting speech act tries to repair; (b) the means available for doing so, depending on the type of goal state to be achieved (belief, desire, emotion, action, etc.). Finally, in section 5 we outline the general idea for how our approach extends to other kinds of relations between speech acts, going beyond support.

2. Relations between speech acts

This section reviews some previous approaches to coherence at the level of speech acts, focusing especially on the question how relations between speech acts make a discourse coherent, and, conversely, what it means for a discourse to be incoherent at speech act level. Section 2.1 looks into the matter from the perspective of the by far better understood coherence relations at propositional level and argues against the common assumption that the same coherence principles apply at both levels. Section 2.2 summarises some relevant insights concerning coherence relations between speech acts across speakers in dialogue. Section 2.3 discusses some previous ideas on how these insights can be made fruitful to understand coherence between utterances of the same speaker, within a single conversational turn or in monologue more generally. Finally, building on these ideas, section 2.4 introduces the notion of *support relations*, a subclass of relations between speech acts whose contribution to discourse coherence is the main subject of this paper.

2.1 Semantic relations at speech act level

Many existing approaches to discourse coherence (Halliday and Hasan, 2014; Pierrehumbert, 1980; Sanders et al., 1992; Sanders, 1997; Sweetser, 1990) make a distinction between coherence relations at propositional and at speech act level. The exact boundary might be drawn slightly differently depending on the approach, but the general idea is roughly the same. Here we will refer to Sanders et al.'s (1992, pp. 7–8) definitions of what they call *semantic* vs. *pragmatic* relations:

> '*A relation is semantic* if the discourse segments are related because of their propositional content. In this case the writer refers to the locutionary meaning of the segments. The coherence exists because the world that is described is perceived as coherent.'
>
> '*A relation is pragmatic* if the discourse segments are related because of the illocutionary meaning of one or both of the segments. In pragmatic relations the coherence relation concerns the speech act status of the segments. The coherence exists because of the writer's goal-oriented communicative acts.'

Note the asymmetry between these definitions. While the definition of semantic relations clearly specifies what it is about the locutionary meaning of the discourse segments that makes the sequence coherent (namely the coherence of the world described), the definition of pragmatic relations does not specify what it is about the goal-oriented communicative acts that creates coherence. However, in practice, the authors simply transfer the same set of relations that are motivated from the point of view coherence in the world (e.g. causal relations) to the domain of speech acts. Although the issue is not explicitly addressed in that work, in our understanding, this basically amounts to saying that coherence between speech acts is governed by the same principles as coherence between propositions and ultimately depends on coherence between states and events in some relevant world. While this is a sensible null hypothesis, here is why we believe it is wrong.

First of all, what does it mean for a world to be coherent? Following David Hume's classification of relations between ideas, Kehler (2002) proposes that there are exactly three ways in which elements of the described reality can cohere or belong together—by causal relations, by spatio-temporal contiguity, or by resemblance (similarities and differences). If we can recognise such relations between objects, states, and events in the world, we perceive that world as coherent, otherwise we don't. Consequently, all coherence relations at propositional level belong to one of the three classes. And if the addressee cannot recognise the described events or states as causally connected, contiguous, or in some relevant sense similar, then he or she will not perceive the discourse as coherent.

Second, what does it mean to apply this idea to speech acts? Since causal, contiguity and resemblance relations can hold between events, and speech acts *are* events, then nothing prevents the same relations from holding between speech acts. That is, speech acts can cause other speech acts, they can be spatio-temporally contiguous to other speech acts, or they can resemble other speech acts. But is this enough to make the sequence of speech acts coherent? Examples like (3) clearly indicate a negative answer: The two assertive speech acts are spatio-temporally contiguous, but the discourse is nevertheless perceived as incoherent.

(3) # John broke his leg. I like plums. (Knott and Dale, 1994)

Moreover, any pair of adjacent speech acts in discourse are spatio-temporally contiguous by definition. If spatio-temporal contiguity were enough to make a discourse coherent at speech act level, then all sequences of speech acts would be coherent. Clearly, such a notion of discourse coherence would not be very useful.

A possible objection to this argument could be that it is not the *actual* spatio-temporal contiguity or adjacency that is relevant to establish coherence between speech acts, but

the *expected* adjacency (see Fetzer, 2013, on a related distinction between adjacency position and adjacency expectation). The sequence of speech acts in (3) is incoherent, because the second speech act presents an unlikely, unexpected follow-up to the first. In contrast, (4) is coherent because the second speech act is a felicitous answer to the question and in that sense is expected in that context. One could even argue that A's question causes B to answer, and therefore a causal coherence relation holds between these speech acts.

(4) A: What's your favourite fruit?
B: I like plums.

The problem with this view is that it does not really give an answer to the question what makes (4) coherent at speech act level, but simply restates the question. Intuitively, (4) is coherent for a different reason than (1-b), repeated in (5), or (2) is. We want to understand that difference rather than just saying that all three are coherent because some speech act presents a likely follow-up to another.

To be fair, there is a strong case for assuming some semantic relations, most notably causality, to hold at speech act level where those relations are signalled by a connective, such as *because* in the classical example from Sweetser (1990):

(5) a. What are you doing tonight?
b. Because there's a good movie on.

Similar speech-act-level uses have been found for many other connectives that are causal in a broad sense, including conditionals and concessives, which led to adopting corresponding coherence relations, for instance, in the annotation scheme of the Penn Discourse Tree Bank (Webber et al., 2019). Typically, the connective establishes a relation between a speech act as its one argument, e.g. (5-a), and a proposition as the other, e.g. (5-b).

However, transferring this approach to cases where there is no explicit connective and both arguments are speech acts is problematic for the reasons stated above. In our own experience of corpus annotation (Jasinskaja et al., 2024), we were often tempted to categorise a relation as speech-act-level causality where we simply could not find a more suitable label. We believe that this is because at a certain level of abstraction, every speech act is in some sense caused by the previous discourse: we give answers *because* we are asked questions, we clarify *because* the previous speech act was unclear, we give arguments *because* the previous speech act was not entirely convincing as it stands. In other words, the definition of speech-act-level causality fits almost any coherent sequence of speech acts. But this is not a very interesting notion of speech act coherence. Instead, our goal in this paper is to give a more specific characterisation of the different ways in which speech acts can cohere.

2.2 Relations between speech acts in dialogue

Interestingly, the foundation for an answer to our main question was already laid by Austin (1962, p. 160), who pointed out that many speech acts have illocutionary forces that "invite by convention a response or sequel". For example, if one asks someone to do something, the hearer is invited to perform a certain action, if someone asks a question the hearer is invited to answer the question and so on. Searle (1969, 1983)

labels the relevant property of a speech act as the *condition of satisfaction*. For example, the condition of satisfaction of an order is that the hearer obeys that order, the condition of satisfaction of a question is that the hearer gives an answer to that question, and so on. Obviously, not all speech act types will be "satisfied" by another speech act, but some will (question–answer pairs being the paradigmatic case), and in those cases one could say that the sequence is coherent because the second speech act fulfils the satisfaction condition of the first.

Within the philosophical tradition of speech act theory, this idea was further developed by Sbisà (2002), who argued that conventional effects of illocutionary acts make sequences of speech acts in general possible, moreover, that sequences of speech acts make it possible that a speech act achieve its conventional effect. However, there have been altogether few attempts to apply the notions of traditional speech act theory to explain coherence of speech act sequences (notable exceptions being Sbisà (2002); Franke (1990), Fritz and Hundsnurscher (2009) who analysed possible reactions to accusations and Hindelang (2010), who classified speech acts in sequences based on the work of Franke, Fritz and Hundschnurscher).

The issue received more interest in conversation analysis. In particular, Sacks and Schegloff coined the term *adjacency pair*, referring to pairs of utterances where the second one constitutes a proper reaction to the first one. Typical examples of such pairs are *question – answer*, *greeting – counter greeting* or *offer – acceptance/refusal*. Sacks and Schegloff (1973) did not attempt to explain these pairings in terms of the notions of speech act theory, but what is clear from the known listings is that different illocutionary forces are paired with different types of reactions. Within this approach one could say that a pair of speech acts is coherent if it constitutes an adjacency pair, and even if we do not fully understand which properties of the speech acts are responsible for the legitimate pairings, there clearly is a connection.

Finally, an utterance may constitute a coherent follow-up to another one, not only if it satisfies the expectations set up by the first utterance, but also when it frustrates those expectations, but only in a way that ultimately serves the purpose of the conversation. Franke (1990), for example, introduced *decision-preparing reaction moves*, which are performed to help decide between a positive and a negative response to a previous speech act. Franke's decision-preparing moves include clarification requests (which were also studied at length in computationally oriented approaches to dialogue, e.g. Ginzburg, 2015) and speech acts that raise a problem related to a previous speech act. With respect to speech act level coherence that means in particular that, for instance, a clarification request can constitute a coherent follow up to a speech act of any type, as long as it helps solving communicative problems related to the first utterance.

In summary, a pair of speech acts can be coherent either because the second speech act satisfies the goals or expectations set up by the first, or because it signals that the first speech act failed to achieve its original goal. In all the cases reviewed so far, however, the first and the second speech act are produced by different speakers, so one might get the impression that coherence at speech act level is a phenomenon restricted to dialogue interaction. The next section shows what we can learn from dialogue about speech act level coherence in monologue.

2.3 From dialogue to monologue

Since dialogue involves a constant exchange of roles between two speakers S1 and S2, and after S2's reaction the turn usually comes back to S1, one might wonder whether the relationship between S1's initial speech act and her reaction to S2's reaction can be fruitfully analysed in speech-act-theoretic terms. For instance, S1 could use an ambiguous pronoun in the first move (6-a), S2 could react with a clarification request in the second move (6-b), to which S1 could provide the desired clarification in the third move (6-c). The question is then: How do the first and the third utterance, (6-a) and (6-c), cohere at speech act level? Which properties of these utterances as goal-oriented communicative acts are essential for the overall coherence of the sequence?

(6) a. $S1$: At that point, it was over across the road.
b. $S2$: What do you mean 'it'?
c. $S1$: The warehouse.

A systematic classification of 'speech acts of the third move' has been developed by Franke (1990). After a clarification request or some other kind of reactive move that frustrates the goals of the initial utterance, the speech act in the third move may belong to one of three classes according to the relationship between its goals and the goals of the initial speech act: *Retractive speech acts* signal that S1 completely gives up on the goal of her initial speech act; *revising speech acts* modify the communicative goal of the initial speech act in such a way that it is more likely to be achieved; finally, *re-initiative speech acts* present a second attempt to achieve the goals of the initial speech act without modification. The third move in (6) is an instance of a re-initiative speech act. Just as in (6-a) S1 is trying to inform S2 that the warehouse was across the road, in (6-c) she is still trying to do so. In this respect, (6-c) presents a coherent follow-up to (6-a) and (6-b) because it continues to pursue the goals of (6-a) after (6-b) indicates that they were not reached in the first attempt.

Notice that 're-initiation' (using Franke's terminology) is a relation between utterances of the same speaker. But speakers often do not wait for an explicit clarification request from their interlocutor, but detect potential communicative problems by monitoring their own speech, and produce re-initiative speech acts, more commonly known under the label of *self-repairs* (Levelt, 1983), without conversational turn transition. In this way, 're-initiation' can function as a speech act level coherence relation in a monologue. This idea is developed by Ginzburg et al. (2014) as a general approach to self-repair moves and other kinds of speech disfluencies. In this regard, (7) has the same underlying structure as (6), with the difference that in (6) the clarification request is made explicit, whereas in (7) it constitutes an implicit question under discussion (QUD).[2]

(7) At that point, it, the warehouse was over across the road.

As far as the source of coherence is concerned, we could say that *the warehouse* constitutes a coherent interjection within the utterance *At that point, it was over across the road* because it ultimately helps achieve the initial goal of informing the hearer that the warehouse was across the road. Coherence is established at the level of the communicative goals.

2. (7) is the original example cited by Ginzburg et al. (2014). The example in (6) is constructed from it.

2.4 Support relations

Adjacency pairs, clarification requests, retractions, revisions, re-initiations, and (self-) repairs are relational categories developed in previous research that, in our view, genuinely pertain to the speech act level of coherence because the source of coherence (the reason why a sequence is coherent) in all those cases has something to do with the communicative goals of the speech acts involved and the relative success or failure in achieving those goals. Of course, causality also plays an important role here since one speaker communicating a certain goal can *cause* the other speaker to want to satisfy that goal, or the failure to achieve a certain goal can *cause* the speaker to start another attempt, but as we argued in section 2.1, in order to achieve a deeper understanding of coherence at speech act level, we should shift our focus from causality in general to more specific relationships between communicative goals. Therefore, the focus of this paper will be on a subclass of relations between speech acts that we will call *support relations* and define as follows:

(8) SUPPORT RELATIONS:
Speech act S of speaker A *supports* speech act N of the same speaker iff
   a. A believes that N has failed or will fail to achieve its goal
   b. A believes that S will help to achieve that goal.

As should be clear from the discussion in previous sections, this definition primarily covers self-repairs and Franke's re-initiations. It is also closely related to the notion of *presentational relations* (9) in RST (Mann and Thompson, 1988):

(9) PRESENTATIONAL RELATIONS: Presentational relations are those whose intended effect is to *increase some inclination* in the reader, such as the desire to act or the degree of positive regard for, belief in, or acceptance of the nucleus. (Mann and Thompson, 1988)

All presentational relations in RST are relations between a nucleus (N), a discourse unit that is more central to the overall purpose of the discourse, and a satellite (S), a discourse unit whose function is defined in relation to the nucleus. The N and S notation in (8) is our adaptation of this distinction to support relations, where S stands for support, and N for the unit being supported.

Although Mann and Thompson's definition does not explicitly mention speech act goals, there is a substantial overlap in the range of cases covered by the definitions (8) and (9). Both types of relations have in common that S has the function to increase the chances of the success of N. For instance, if the goal of N is that the hearer believe a certain proposition and S increases the hearer's belief in that proposition, as in the case of *Evidence* relations like that in (10), then S also helps achieve the goal of N.

(10) a. He must have been here recently.
   b. There are his footprints.

However, in this paper, we pursue a methodological approach radically different from that of RST. The approach of RST is empirically driven and bottom-up, the proposed set of rhetorical relations is motivated by what can be found in naturally occurring texts, and is potentially open for new additions. This approach has been criticised for its

failure to give a principled basis for saying what a theoretically (im)possible coherence relation is, since its descriptive nature only allows to account for relations that can be found in existing texts (Knott and Dale, 1994). In this paper, we address that challenge by taking a strictly top-down approach. We first try to give answers to the following questions: What kinds of goals can speech acts have? What are the possible ways to fail those goals? What are the possible ways to achieve those goals? That provides the basis for our taxonomy of all theoretically possible support relations. As will become clear, this approach does not only reveal systematic similarities between self-repairs, Franke's re-initiations and RST's presentational relations, but also creates previously unknown categories for speech act level coherence relations.

One last remark before we embark on this endeavour: Presentational relations in RST do not impose any general constraints on the linear order of the nucleus and the satellite, both (N, S) and (S, N) sequences are possible. Repair moves can also generally follow the reparandum, or be linearly embedded in it, as in (7) (or even precede it, although in this case the reparandum is rarely fully articulated, cf. Forward Looking Disfluencies in Ginzburg et al., 2014). In principle, supporting speech acts can also follow, precede or be linearly embedded in their nuclei. However, this paper will mainly focus on the case of (N, S), where support follows the nucleus.

3. The goals of a speech act

As defined in the previous section, supporting speech acts and support relations are produced in order to help achieve the goal(s) of a speech act which is (anticipated to be) unsuccessful. Given this goal-based characterisation of a supporting speech act, one should wonder which goals can be achieved by a speech act and therefore, what can go wrong.

For a speech act to be fully successful, a whole cascade of goals have to be achieved. The intermediate steps with their respective goals that we assume to lead to communicative success (inspired by Clark's 1994; 1996, levels of action in communication) are listed in figure 1. For an illustration, let us consider the speech act *Off with their heads!* performed by the Queen of Hearts upon seeing that her gardeners had planted roses in the wrong colour in Lewis Carroll's novel *Alice's Adventures in Wonderland*:

(11) 'I see!' said the Queen, who had meanwhile been examining the roses. 'Off with their heads!' and the procession moved on, three of the soldiers remaining behind to execute the unfortunate gardeners, who ran to Alice for protection.

In spoken communication, the first step is the production of a certain acoustic signal. The goal is to get sound waves of appropriate shape to travel to the right place at the right time (signal transmission). Second, the acoustic signal must be processed by the hearer's auditory system, resulting in a certain pattern of activation in the hearer's brain and, presumably, a representation in the hearer's mind (signal processing).

The next steps correspond more or less closely to Clark's levels of action in communication.[3] After the sounds have been recognised as speech, the hearer must first of all pay attention to the signal (*attention*), which is a prerequisite for all further pro-

3. Clark emphasises the joint character of communicative action in which both the speaker and the hearer have a role to play. He identifies four levels of action: 1. vocalisation and attention, 2. presentation and identification 3. meaning and understanding and 4. proposal and uptake (the first term

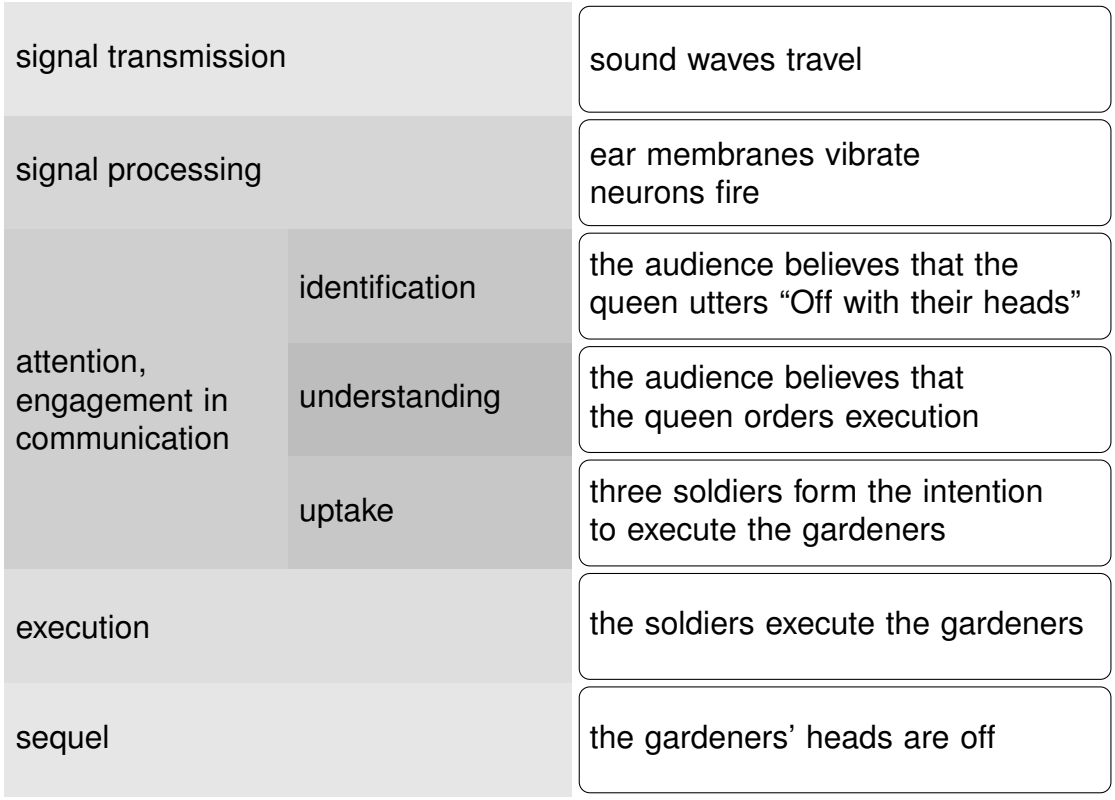

Figure 1: The *workflow* of the speech act *Off with their heads!* in (11).

cessing. While Clark treats attention as the first step in the sequence on a par with the subsequent steps, we see it as a state that must hold for the whole duration of communication and in that sense cannot be meaningfully ordered with respect to the other steps.

Provided that the hearer is paying attention and engaging in the communication, they must identify what the speaker *said* and what the speaker *meant*. In the next step, the queen's audience has to map the sounds they hear to the words *Off, with, their* and *heads*, i.e. create a representation of the linguistic form of the utterance (*identification*). Then, they must come to the belief that the queen produced these words because she wanted to order the execution of the gardeners. That includes recognition of both the semantic content of the utterance and the speaker's communicative intention (*understanding*). After the queen's message was understood, the ultimate success of the speech act will depend on whether or not the hearer reacts in the expected way. We divide the hearer's reaction into a 'mental' part (*uptake*) and a 'physical' or 'active' part (*execution*). The 'mental' part of the goal in the present example is that some of the soldiers (three out of ten taking part in the procession) form the intention to execute the gardeners.

A few clarifying remarks on our notion of uptake are in order. The term dates back to Austin (1962), who used it in a rather broad sense to refer to the whole range of effects a speech act can have upon the hearer, including Clark's identification, under-

in each pair describes the speaker's part and the second term the hearer's part in the joint action). In this paper, we think of the hearer's part as the goal to be achieved by the speaker's action.

standing, uptake in the narrow sense, and probably our execution as well. Clark (1996) draws a clear line between uptake and understanding: Understanding is the correct construal of the speaker's action, e.g. an order to behead the gardeners is recognised as an order to behead the gardeners. Uptake, in turn, is the hearer's action upon the proposal. However, Clark seems to include a number of rather different things in that.

On the one hand, Clark's notion of uptake includes appropriate reactions in dialogue, i.e. roughly, the second parts of adjacency pairs. For instance, an answer to a question constitutes the hearer's uptake of the question. To the extent that the notion of adjacency pair is taken to cover cases of non-speech action, the soldiers' action of beheading the gardeners in our example constitutes the uptake of the Queen's order (see also Hulstijn and Maudet, 2006).

On the other hand, Clark also counts mere consideration of the speaker's proposal by the hearer as uptake. In our present example it would mean that uptake already takes place when the soldiers just contemplate whether or not they should follow the Queen's order. This weaker notion of uptake has been more widely adopted in subsequent research, to the point of complete replacement of the notion uptake by the notion of consideration (see e.g. Rodríguez and Schlangen, 2004).

In the very narrow sense of uptake that we would like to adopt in this paper, neither counts as uptake. Consideration of the speaker's contribution, in our view, constitutes part of paying attention to the exchange and engagement in communication. As pointed out above, we assume that attention and engagement must be given at all levels of utterance processing, and consideration of the proposal should therefore be seen as engagement in communication at higher levels (at the level of uptake, in particular). Appropriate reactions in dialogue, such as answering a question or fulfilling an order, on the other hand, are *executions* of actions that may serve as evidence of the hearer's uptake, but they do not make up its essence.

In this paper, we will use the term 'uptake' to refer to the hearer's mental reaction to the speaker's utterance that is in accordance with the speaker's desires and intentions. This notion comes close to what Schlöder and Fernández (2015) call *intention adoption* (reaching mutual agreement), which they distinguish from *intention recognition* (understanding that goes beyond semantic interpretation). Along the same lines, we see intention recognition as part of understanding the utterance, which captures the pragmatic rather than semantic aspects of understanding, whereas the adoption of the speaker's intention is the hearer's mental reaction intended by the speaker.

What kind of mental reaction that is will generally depend on the speech act type. Uptake of a directive speech act, such as the Queen's order, is the adoption of the intention to fulfil that order. Uptake of an assertion is usually the belief of the asserted proposition. Uptake of an insult is feeling insulted, and so on. In any case, uptake goes beyond mere understanding of an utterance and includes mental compliance with the speaker's proposal, but does not go as far as the (physical) action that might result from that mental compliance, such as actually carrying out the order or signalling by means of a nod or a feedback utterance (e.g. *yes*, *mhm*) that the hearer accepts the asserted proposition. The latter belong to the level of execution.[4]

4. Admittedly, in some cases it might be difficult to draw a line between uptake and execution, especially in directives that concern mental acts. For instance (i) is literally a request to make the stated assumption, but it is hard to imagine forming an intention to make this assumption without making

Finally, the goal of a speech act may go beyond the hearer's mental attitudes or physical actions and lie anywhere in the outside world. Any direct or indirect consequences of the speech act may constitute its ultimate intended effect. For instance, the state of affairs of the gardeners' heads being 'off' could be seen as the goal of the Queen's order at the level of the *sequel*. Note that the Queen does not specify how the gardeners' heads should get 'off', or who should do what to achieve that result. In fact, as we know from the further development of the story, Alice ensures that the gardeners' heads are 'off' without being parted from their bodies, and everyone appears happy with that solution. We could say that the *literal goal* of this speech act is that the gardeners' heads are 'off'. This is also the *ultimate goal* of this speech act. That is, both the literal and the ultimate goal are located at the level of the sequel.[5]

Speech acts of different types expressed by different sentential moods will generally have literal/ultimate goals located at different levels. For instance, a typical directive expressed by an imperative sentence, such as (12), literally expresses the goal that the hearer perform the action of leaving the room. That is, the literal goal of this speech act (and taken at face value, also its ultimate goal) lies at the level of execution.

(12) Leave the room.

One might disagree on whether the goal that the hearer believe that it is raining is literally expressed by (13) or is the result of pragmatic inference. However, that goal clearly lies in the hearer's mind and therefore at the level of uptake. This is also its ultimate goal if the speech act is taken at face value, but if it is interpreted as an indirect speech act, e.g. as advice to put on appropriate clothing, then the ultimate goal lies at the level of execution.

(13) It is raining.

In other words, the ultimate goal of a speech act may pertain to the levels of uptake, execution or sequel, but reaching that goal will require reaching intermediate goals at all the other levels leading up to it, which includes attention, identification and understanding. The linguistic form of the utterance, in particular its sentential mood, gives the hearer a clue of the intended ultimate goal, though it does not necessarily literally express it. Things can go wrong at any of these levels. In the next section, we develop a taxonomy of support relations based on the location of the communicative problem as the main criterion for classification.

---

the assumption yet. It is not theoretically impossible, but something we probably very rarely do with simple mental actions of this kind.

(i) Suppose x is greater than 0.

5. To put these levels in relation to the standard speech-act-theoretic notions, reaching a goal at the level of identification corresponds roughly to performing the phonetic act in the sense of Austin (1962, p. 95). Reaching semantic understanding corresponds to the performance of a *locutionary act*, whereas reaching pragmatic understanding or what Schlöder and Fernández (2015) call intention recognition corresponds to the performance of an *illocutionary act*. Finally, achieving uptake, execution, or any more far-reaching goals that pertain to the sequel belongs in the domain of *perlocution*.

4. Towards a taxonomy

As we promised in section 2, the primary motivation for the taxonomy of support relations to be developed in this paper is theoretical rather than empirical. Our goal is to circumscribe all theoretically possible kinds of support relations based on the concept of support as a relation between speech acts. The classification criteria should therefore be based on the understanding of what support is and what a speech act is, on the constitutive parts of these notions and properties that are associated with them by definition.

The definition of support relations given in (8) is repeated below. Two main criteria for the classification of support relations follow from the two conditions in this definition. One can distinguish between different kinds of support depending on (a) how N failed, and (b) how S solves that problem.

(14) SUPPORT RELATIONS:
Speech act S of speaker A *supports* speech act N of the same speaker iff
   a. A believes that N has failed or will fail to achieve its goal, and
   b. A believes that S will help to achieve that goal.

In section 3, we offered our version of the levels of action that expands on and modifies that proposed by Clark (1996), cf. figure 1. However, the basic observation concerning these levels remains the same: a speech act can fail at any of the levels that it involves, and therefore a supporting speech act may be called for to repair a problem at any of these levels. So, the first dimension of our taxonomy is the location of the problem targeted by the supporting speech act. Section 4.1 gives an overview of the types of support relations according to this criterion. While we will have less to say about the second dimension—the types of support relations according to how they solve the problem at hand—section 4.2 outlines our general strategy to approach this issue and showcases one group of support types aimed at affecting the hearer's beliefs.

4.1 Locating the problem

SUPPORT FOR ATTENTION / ENGAGEMENT IN COMMUNICATION

A classical example of a speech act that supports another speech act because the latter failed at the most basic level of securing the addressee's attention is (15) from Clark (1994). Here it seems that Bob did not even hear Ann's initial utterance (15-a). Ann solves the problem by repeating it more loudly (15-c). That is, (15-c) supports (15-a).

(15) a. Ann: Bob
   b. Bob: [3 sec of no response]
   c. Ann: Bob [louder]
   d. Bob: What?

Clark's original ladder of levels of action in communication creates the impression that attention, as the first step of that ladder, is on a par with the other steps in the sense that success at higher levels presupposes success at the attention level. This, in turn, seems to be based on the implicit assumption that once attention is achieved it cannot be taken away, similar to how understanding, once achieved, cannot be taken away.

However, as we pointed out in section 3, while identification, understanding and uptake are events where one is a precondition for the next one, attention is a state that needs to be given throughout the whole process of communication. Research on the psychology of attention has established that while attention can be induced as an automatic reaction to abrupt changes in the environment, it is generally given and maintained intentionally and is driven by the agent's domain-level goals and interests (Yantis and Jonides, 1984, 1990; Theeuwes, 1991; van der Lubbe and Postma, 2005).

Thus, the initial 'grabbing' of the addressee's attention that we see in (15) only marks the beginning of the attention state, after which attention needs to be maintained intentionally by the addressee. Example (16) shows that attention, or more generally the willingness to engage with the addressee, can be taken away after the utterance is identified and understood: (16-e) makes clear that the first person narrator (the addressee) has perfectly understood 'the lady's' attempt at contact. Nevertheless, he denies her his attention and refuses to engage in the communication. It does not matter whether it is obvious to the speaker that the addressee has understood her utterances. By uttering (16-f) and (16-g) she is trying to (re)gain the addressee's attention. Therefore, (16-f) and (16-g) stand in a support relation to all or any of (16-a), (16-c) and (16-d).

(16)
a. "Sorry,"
b. I hear somebody next to me say.
c. "Aren't you the man from the television?
d. The one who harassed those poor girls?"
e. Fuck... I don't answer [...]
f. "Hey,
g. I am talking to you."
h. The lady insists again.[6]

Support for identification

This category of support relations mainly consists of self-repairs that target the hearer's failure to create the correct representation of the linguistic form of an utterance. In Clark's (1994) example, (17-c) supports (17-a).

(17)
a. A: yes forty-nine Skipton Place
b. B: forty-one
c. A: n i n e . nine
d. B: forty-nine, Skipton Place,

Support for understanding

In section 3, we adopted a broad notion of understanding, which includes both semantic and pragmatic understanding. An utterance counts as fully understood only if the hearer is able to correctly identify the speaker's meaning behind it (Grice, 1957). That includes understanding the conventional meanings of the words and phrases, reference resolution, presupposition resolution, being able to draw the intended implicatures, understanding what kind of speech act the speaker intends to perform by means

6. From *Almost* by Adriana LS Swift.

of the utterance, understanding whether the utterance is meant seriously or jokingly, etc.

Support relations that target semantic aspects of understanding are otherwise known as self-repair and reformulation. For instance, the self-repair in (18-c) clarifies the sense in which Ken uses the verb *evaluate* in (18-a).

(18) a. Ken: k who evaluates the property - - -
b. Ned: uh whoever you asked, . the surveyor for the building society
c. Ken: no, I meant who decides what price it'll go on the market -
d. Ned: (- snorts) . whatever people will pay

In (19), another example from Clark (1994), Sam's response *m* in (19-c) to Dar's clarification request supports (19-a) by confirming the reference of *this boy*:

(19) a. Sam: well wo uh what shall we do about uh this boy then
b. Dar: Duveen
c. Sam: m

These are instances of more or less spontaneous self-repair. But speech acts whose purpose is to improve the understanding of a previous utterance can also be planned. In (20), the speaker does not only replace the (presumably) unfamiliar term *anacrusis* with a (presumably) more accessible definition, but the reformulation is used as a strategy to establish the equivalence of *anacrusis* and *an unaccented note which is not part of the first full bar* (Blakemore, 1993), i.e. to define the new term.

(20) a. This piece begins with an anacrusis,
b. an unaccented note which is not part of the first full bar.

Mann and Thompson's RST (Mann and Thompson, 1988, p. 273) includes a presentational relation *Background*, whose definition covers the essence of a support relation that targets a potential problem at the level of understanding (R: reader; N: nucleus; S: satellite):

(21) *Background*:
a. R won't comprehend N sufficiently before reading text of S
b. S increases the ability of R to comprehend an element in N

However, the cases that this definition is normally applied to in RST are different from those mentioned above. First of all, in RST practice, the typical order of discourse units connected by a *Background* relation is (S, N), rather than (N, S). In (22) from Mann et al. (1989), the first sentence presents the event of media covering the results of ZPG's 1985 Urban Stress Test, which satisfies the existence presupposition of the definite NP *this remarkable media coverage* in the second sentence. Without (22-a) preceding it, (22-b) as it stands would be unacceptable in formally published written discourse, although with slight modifications the reverse order would also be possible, cf. (23).

(22) a. The results of ZPG's 1985 Urban Stress Test were reported as a top news story by hundreds of newspapers and TV and radio stations from coast to coast.

b. I hope you'll help us monitor *this remarkable media coverage* by completing the enclosed reply form.

(23) a. I hope you'll help us monitor *the remarkable media coverage of the results of ZPG's 1985 Urban Stress Test* by completing the enclosed reply form.
b. The results were reported as a top news story by hundreds of newspapers and TV and radio stations from coast to coast.

The second sentence in (23) reads like an afterthought that one would like to put in parentheses, and in that sense resembles the more spontaneous self-repairs in (18) and (19), which is probably why *Backgrounds* that follow their nuclei are less acceptable in written texts. However, what is common to (22) and (19) is that in both cases the purpose of the supporting speech act is to provide information that helps establish the reference of a definite description, *this remarkable media coverage* and *this boy*, respectively.

Support relations that target pragmatic aspects of understanding have received less attention in previous research, or might have partly been handled under different unrelated categories. Understanding the speaker's intention behind the utterance involves many different layers, including the understanding of implicatures, illocutionary force, perlocutionary object, among others. Example (24) is an instance of conversational implicature clarification. The first sentence (24-a) has two readings: the 'normal' one, without any notable quantity implicature, and the one where the quantity implicature *like* $\rightsquigarrow$ *not love* is drawn in the scope of negation. The second sentence (24-b) makes it clear that the more marked second reading is intended.

(24) a. Around here, we don't like coffee, (Horn, 1989, p. 382)
b. we love it.

In (25), an excerpt from the novella *The Little Prince* by Antoine de Saint-Exupery, the Little Prince is in conversation with the King, who he meets on one of the planets he visits on his journey. In (25-d), the King orders the Little Prince to yawn. Just in case the exact illocutionary force of the imperative might be misunderstood, (25-e) supports (25-d) by clarifying that it is an order.

(25) a. It is years since I have seen anyone yawning.
b. Yawns, to me, are objects of curiosity.
c. Come, now!
d. Yawn again!
e. It is an order.

A supporting speech act may be called for if the perlocutionary goal of the nucleus needs to be clarified. For instance, (26-a) might be taken as an insult. In (26-b), the speaker makes clear that he does not intend to insult the hearer.

(26) a. Fuckin' I hate this guitar! I hate it so much!
b. No offence, Ed.[7]

7. From an interaction between Thom Yorke and his band during a studio recording: https://www.youtube.com/watch?v=6eRp97ZRwmk, time stamp 4:45–4:50. Last accessed August 24, 2023.

The following example stems from a reference manual for Ableton Live 5, digital audio workstation software.[8] The purpose of user manuals is to give instructions about how to use a product. Instruction is a subtype of a directive speech act that is appropriate to perform in a situation where the addressee has a certain goal in mind which he or she desires to achieve, but does not know how to do it. Therefore, by saying (27-a), the speaker insinuates that the addressee might want, among other things, to get rid of unwelcome house guests or terrifying small pets. In the supporting speech act (27-b), the speaker corrects for a potential misunderstanding, stating that (27-a) was meant as a joke.

(27) a. This can be very useful in creating new sounds and textures, as well as getting rid of unwelcome house guests, or terrifying small pets
b. (just kidding!).

Sequences like (25-d)–(25-e), (26-a)–(26-b), (27-a)–(27-b) do not seem to fit comfortably into the existing taxonomies of coherence relations. While they might literally fit the RST definition of *Background*, they do not fit the intuitive notion of background as something that *precedes* something in some sense (temporally, epistemically, or in the flow of communication) and they all show the non-canonical (N, S) order of segments. None of the PDTB speech act relations seems to apply in these cases.

Finally, there are supporting speech acts whose purpose is to establish the relevance of another speech act. Sweetser's (1990) classical examples of speech-act causality (28), as well as some instances of the RST relation *Justify* (29) belong to this category.

(28) a. What are you doing tonight?
b. Because there's a good movie on.

(29) *Justify*:
R's comprehending S increases R's readiness to accept W's right to present N

The question in (28-a) might be understood as asking for information the speaker is not entitled to. However, (28-b) explains why the question is relevant. By uttering (28-b) the speaker shows that they basically mean it as an invitation to go to see the film together, something that is meant for the addressee's benefit as much as their own. After understanding (28-b), the addressee does not even need to give a full answer to the original question. It is enough if they say whether or not they will go to the movies with the speaker as that is ultimately the only thing that is relevant.

SUPPORT FOR UPTAKE

The notion of uptake adopted in section 3 refers to the hearer's mental reaction that goes beyond mere understanding of the speaker's intention and encompasses cooperative adoption of that intention. It is the success of the utterance at the perlocutionary level, albeit limited to its mental component. The type of mental reaction in question will depend on the goal. For instance, the perlocutionary goal of an assertion is often to make the speaker believe a proposition. Therefore, the uptake of the assertion would consist in the hearer's adopting that belief. The perlocutionary goal of a directive is

8. URL: `http://downloads.ableton.com/manuals/50/ableton_live_5_manual_en.pdf`. Last accessed August 24, 2023.

to bring the hearer to perform an action. The uptake of the directive would be the hearer's forming an intention to perform that action. Etc.

Support relations whose purpose is to secure uptake can be distinguished according to the type of mental act or state of the hearer that constitutes the goal of the utterance. If the goal is the addressee's belief in a certain proposition, then the supporting speech act should make that proposition more believable. The RST relation *Evidence* is geared towards this situation:

(30) *Evidence*
   a. R might not believe N to a degree satisfactory to W
   b. R's comprehending S increases R's belief of N

A typical *Evidence* is an assertion that states something more evident, i.e. something more directly observable, than the proposition expressed by the nucleus:

(31) a. He must have been here recently.
   b. There are his footprints.

But giving *Evidence* is not the only way to make a proposition more believable. In section 4.2, which discusses distinctions between support relations according to how the communicative problem is *solved*, we will present a number of other support relations that also target the addressee's belief.

If the goal of the utterance is ultimately to bring the addressee to perform a certain action, the mental state that normally precedes that is the addressee's desire or intention to perform that action. This is typical for directive speech acts. If a directive is anticipated to fail, the speaker might perform a supporting speech act to make the addressee more willing to perform that action. The RST relation *Motivation* (32) covers this case neatly. For instance, (33-c) provides a reason why the google development team might want to fulfil the request expressed in (33-b).

(32) *Motivation*
   a. N presents an action in which R is the actor, unrealized with respect to the context of N
   b. comprehending S increases R's desire to perform action presented in N

(33) a. I have an idea,
   b. why don't you guys add the pause and download on the playstore,
   c. that way it would increase your downloads[9]

Belief and desire/intention are the first kinds of uptake that come to mind as they are directly associated with two kinds of speech acts—assertions and directives.[10] Speech

9. A message posted August 7, 2023 on the Google Play Help forum: https://support.google.com/googleplay/thread/229269275/i-have-an-idea-why-don-t-you-guys-add-the-pause-and-download-on-the-playstore-it-will-help-you-guys. Last accessed: August 27, 2023.

10. Belief can also constitute the perlocutionary goal of a commissive speech act. If commissives were only about regulating the speaker's own commitments, then what is the point of saying it? Just do, what you are committed to do! By making explicit promises we often try to influence the behaviour of others, for instance, to obtain something in return. That requires that the addressees trust our promises, i.e. believe the propositions they express.

acts can also pursue the goal of triggering emotions, which may or may not require the intermediate step of imparting a belief. For instance, the use of rude vocabulary can be insulting by itself, regardless of whether the addressee believes the proposition expressed by the utterance, or whether the utterance even expresses a proposition. Perlocutionary acts whose purpose is to elicit emotion are altogether less well studied, even less so are the speakers' discourse strategies when those acts fail. Support relations for emotional uptake will have to remain a question for future research. However, it is important to emphasise that they *must* exist because emotional perlocutionary acts exist, and therefore, a comprehensive taxonomy of support relations must include them.

Support for execution

Most speech acts do not require from the addressee the execution of any physical or mental action beyond adopting a certain belief, but those that do—directive speech acts, first and foremost—can also run into problems at the execution level. In (34), the uptake goes smoothly, speaker B readily agrees to perform the action requested in (34-a). However, in (34-c) speaker A provides information that will make it possible, or easier for B to perform that action. In RST, such support relations are called *Enablement*, (35).

(34) a. A: Please can you post these letters?
b. B: Sure.
c. A: The stamps are in the drawer.

(35) *Enablement*
a. N presents an action in which R is the actor, unrealized with respect to the context of N
b. R comprehending S increases R's potential ability to perform the action presented in N

Concluding remarks

We have surveyed different types of support relations according to the location of the communicative problem that the supporting speech act is designed to solve. As it appears, some of them correspond neatly to relations previously defined in the literature (*Evidence*, *Motivation*, *Enablement*), while others cross-cut known categories or are difficult to sort under any categories proposed in previous research.

We took a closer look at support relations at the levels of attention, identification, understanding, uptake, and execution, but did not cover signal transmission, signal processing, and sequel in the proposed ladder of levels of action in communication in figure 1. However, support relations should exist at those levels, too. For signal transmission, it is not even that difficult to find natural examples. We are all familiar with the situation in online meetings during the covid pandemic when a participant forgets to switch on their microphone before they start to speak. Once they notice their mistake, they usually switch on the microphone and repeat their utterance(s).

The combined action of switching on the mic and the repetition stands in a support relation to the initial utterance. This is support for signal transmission.[11]

We are not going to try to find examples for the last remaining two levels. Once again, it follows from the definition of support relations and from the proposed levels of action ladder that such relations *must* exist. We will either encounter them sooner or later in naturally occurring discourse, or we will have to offer a principled explanation why these levels constitute an exception. Both tasks go beyond the scope of this paper.

4.2 Solving the problem

Now we turn to the second criterion for the classification of support relations that follows from the definition in (14)—the way in which the supporting speech act tries to solve the problem caused by the nucleus. It will not be possible to give nearly as broad a survey of possible types as that we gave in the previous section for the criterion of the problem location, but we will present some fundamental considerations on which, we believe, the taxonomy of possible solutions should be based, and we will develop a fragment of the taxonomy for only one, albeit prominent, subclass of support relations, those whose purpose is to induce a belief.

General considerations

Obviously, what means are available for solving a problem will strongly depend on the problem at hand and the nature of the state that the agent desires to achieve. So far we have identified five types of (intermediate or ultimate) goals of speech acts that directly concern the state of the addressee: (a) attention/engagement in communication; (b) belief; (c) desire/intention; (d) emotion; and (e) action.[12] The first thing to keep in mind is that eliciting any of these reactions or behaviours in the addressee requires conscious or intuitive domain knowledge about attention, belief, desire and intention, different emotions, and different types of action.

11. Note that we have systematically defined the ways in which a speech act can fail in terms of missing a certain effect on the addressee. Matej Drobňák (p.c.) pointed out that speakers might also perceive their utterances as failing and might feel the need to produce a supporting speech act when they do not comply to a socially defined standard of propriety or correctness, especially in institutional contexts. For instance, a sentence is valid and legally binding only if the judge pronounces it in the courtroom. Suppose the judge pronounces the sentence while standing in the doorway of the courtroom. They might step back inside the courtroom to pronounce the sentence again to ensure that it has legal consequences. Importantly, it does not matter how the audience takes it. They might have completely understood and agreed with the sentence. Nevertheless, there is some goal the speech act has obviously failed to achieve if the judge feels the need to take another try at it.

    Another example might be self-corrections for spelling and grammar in instant messaging communication with WhatsApp or other applications that allow for later editing of sent messages. Again, the addressee might have understood and adopted the intention of the utterance regardless of the incorrect spelling and grammar. Nevertheless, the speaker feels the need to repair the utterance because it does not comply to some ideal standard of correctness.

    We are not sure yet whether these cases could be reduced to failing some less immediate addressee-related goals (e.g. goals in the sequel) or whether they call for a separate category. This is a point where our taxonomy of support relations might still be incomplete.

12. This lists excludes states that lie outside the addressee's mind or immediate control, those pertaining to the levels of signal transmission, signal processing and the sequel. For reasons of space, those levels will not be discussed any further in this paper.

For instance, we seem to intuitively know that abruptly increasing the volume or making an abrupt movement is likely to draw the addressee's attention (Theeuwes, 1991; Yantis and Jonides, 1984, 1990). Therefore, if our initial attempt to draw the addressee's attention did not work, we might want to increase the speech volume, as in example (15) in the previous section, or accompany our speech act with a (more expressed) gesture. When it is about keeping the addressee engaged for a longer period of time, more sophisticated strategies might be necessary.

1Making people want and intend to perform an action is what persuasion is about. For instance, O'Keefe (2006) distinguishes several major strategies of persuasion depending on the addressee's initial state of mind. One strategy targets the addressee's positive regard of the action itself and its outcomes for the addressee. As in (33), getting more downloads is presumably something the addressee wants, so the speaker chooses that as an argument to persuade them to fulfil his or her request. Other strategies target (a) the influence of the action on the addressee's perception by others; (b) the addressee's perceived ability to perform the action when the desire to perform it is a given and the only thing them from turning the desire into a specific intention is the belief that they cannot successfully perform it; and (c) turning a general intention to perform the action into an intention to do it right now. Accordingly, one could distinguish further subtypes of the *Motivation* relation depending on which strategy is used by the speaker.

If contrary to the speaker's intention, a speech act fails to insult, frighten, console, or amuse the addressee, one needs specific knowledge in the domain of interpersonal emotional (dis)regulation to be able to produce a supporting speech act that will help induce the desired emotion. Some of that knowledge might be intuitively available to the majority of neurotypical population, some may require professional training in therapy, advertisement, propaganda, etc. (Thoits, 1996; Ochsner and Gross, 2005; Niven et al., 2009; Reddy, 2012; Zaki and Williams, 2013)

To ensure successful execution of an action, one requires domain knowledge about that kind of action. For instance, posting letters requires domain knowledge about posting letters and, in particular, the fact that letters need stamps. That is what makes (34-c) a good supporting speech act for (34-a).

In other words, there is probably no single unified taxonomy of solutions for all kinds of communicative problems. The criteria for more fine-grained classification of supporting speech acts according to the solution strategy they employ will depend on the nature of the domain knowledge required. Therefore, it will not be possible to give a comprehensive survey of such criteria and types in this paper. However, to illustrate the basic idea of the approach, we will give a brief overview of strategies for inducing belief and the range of support relations that can be distinguished according to the chosen strategy.

SUPPORT FOR BELIEF

Belief plays a role at different levels of the speech act's workflow, cf. figure 1. The hearer must form the correct belief about the intended linguistic form of the utterance (i.e., *identification*), she must form the correct belief about the communicative intent of the speech act (i.e., *understanding*), and so on. In some cases, most typically in assertions, forming a certain belief is the intended ultimate goal of a speech act.

There are two main approaches to the study of beliefs. The normative approach grounded in philosophy attempts to define the *correct* or *proper* ways of managing our belief states (e.g. Feldman, 2000), driven by the goal of holding only true beliefs as well as by other ethical considerations. The empirical approaches in psychology and cognitive science try to answer the question how humans really acquire, store, and change their beliefs (see Porot and Mandelbaum, 2021, for a recent overview). Both trains of research can give us clues as to what speakers might consider as appropriate or effective ways to induce a certain belief in the addressee, and the classification of discourse patterns of support for belief should be shaped accordingly.

Evidence-based vs. pragmatic: The debate about beliefs has identified two main reasons why an agent may hold a particular belief: evidence-based reasons and pragmatic reasons. According to *evidentialism*, it is only permissible to believe a particular proposition if and only if there are a sufficient number of evidence-based reasons for the truth of that proposition. As Clifford (1877) put it in his seminal paper on the ethics of belief, "it is always, everywhere, and for everyone, wrong to believe anything on insufficient evidence".

The philosophical counterpoint to this view is the assumption that there are also *pragmatic* reasons for believing something. For example, one might believe a proposition because of the benefits of believing it. For example, people might choose to believe that climate change is not anthropogenic because this belief does not force them to think about the consequences of their choices, which is a more comfortable attitude than thinking about one's own responsibility and possible past mistakes. Jordan (1996) argues that there are cases in which it is morally and rationally permissible to form beliefs on the basis of pragmatic reasons.

From a psychological point of view, Porot and Mandelbaum (2021) point out that in some cases belief updating is governed by a 'psychological immune system' that makes it easier to believe propositions that are consistent with our self-image, and to reject beliefs that contradict our self-image. In this sense, the psychological immune system helps the development of pragmatic beliefs.

We expect this dichotomy to be reflected in the strategies to induce belief employed in supporting speech acts. We have already seen an example of evidence-based belief (31) in section 4.1, repeated in (36). Here the speaker wants to induce belief by providing evidence:

(36) a. He must have been here recently.
b. There are his footprints.

On the other hand, supporting moves can offer pragmatic reasons for adopting a particular belief. For example, the speaker of (37) chooses to believe in a better future just because if they stop believing it they might just as well kill themselves.

(37) a. It will get better.
b. Otherwise, what's the point?

A speech act like (37-b) surely also contributes to the psychological immune system, because it reinforces the fact that one's actions are not pointless.

Providing vs. asking to seek: Interestingly, Feldman (2000) points out that our talk about beliefs mirrors our talk about actions, that is, our talk about moral judgement. We say that people should do certain things in certain situations, and we blame them when they commit certain actions. On the other hand, we say that, given a certain amount of evidence, people should believe something, and we blame them if they don't. Accordingly, Hall and Johnson (1998) argue that agents not only have a moral duty to do the right thing, but also an epistemic duty to seek evidence for uncertain propositions. This thesis is called the *epistemic duty thesis*.

Applied to a cooperative communicative situation, this thesis implies that communication partners have a duty to seek evidence for the truth of the proposition in question. In the case where one of the interlocutors wants the other to believe something, the joint duty can be fulfilled by *providing* the missing evidence or by *asking* the hearer *to seek* for evidence.

The first case is exemplified in (36). The second case can occur, for instance, where the speaker refers the hearer to another, perhaps more trustworthy epistemic authority (Zagzebski, 2012), as in (2), repeated in (38). Here the speaker urges the hearer to seek evidence from another source.

(38) a. "Pa told Peter he wanted *him* to be chairman."
b. "Sure he did.
c. If you don't believe me, ask Peter."

Perception vs. testimony vs. inference: Epistemologists have argued that we form our beliefs based on sources of evidence that can be broadly categorised as perception (what we directly perceive), testimony (what other people tell us), or inference (what we conclude based on other evidence), as summarised by Lesage et al. (2015, see also Millar 2011 and Davies and Matheson 2012). Competent speakers are intuitively aware of this distinction, as evidenced by distinct evidentiality marking available in many languages of the world (see e.g. Faller, 2002). So it stands to reason that speech acts of support for belief would make recourse to these three sources of evidence and could be categorised along this dimension.

The standard *Evidence* relation in (36) involves the speaker *telling* the addressee the evidence. Thus for the addressee, the evidence comes from testimony. Similarly, in (38) the speaker asks the addressee to seek further testimony of the proposition expressed by the nucleus.

In contrast, the speaker of (39) *shows* the evidence. By singing the iconic tenor piece from Verdi's opera *Rigoletto* the speaker demonstrates the range of his voice, which the addressee can directly perceive. Contrary to what the RST definition of *Evidence* (30) would require, there is no need to comprehend the supporting speech act. In fact, the addressee need not have any knowledge of Italian to be convinced by the "argument" in (39-b).

(39) a. I'm a tenor.
b. [sings:] *La donna è mobile qual piuma al vento...*

Arguably, in the majority of cases where the target of support is a belief at one of the lower levels in the speech act workflow, cf. figure 1, direct perception will be the most natural way of acquiring evidence by the addressee. For instance, identification of

the utterance is essentially the result of processing the auditory and visual input. If the addressee fails to form the belief that the speaker uttered *nine* as part of (17-a), the most straightforward way to solve the problem is to simply re-enact the same utterance, whose direct perception will help the addressee form the correct belief.

Showing often requires actions that go beyond speech. For instance, performing a double toe loop is a way to show evidence for the statement *I can perform a double toe loop*. If the proposition to be shown concerns facts external to the speaker, showing will typically require drawing the addressees attention to events and states that serve as evidence to those facts. Pointing gestures or presenting pictures, while not strictly *speech* acts, are communicative acts that have this function in multimodal communication.

Rather than providing, the speaker can also ask the addressee to seek direct perceptual evidence. In (40-f), from *Harry Potter and the Goblet of Fire* by J. K. Rowling, Draco Malfoy asks Professor Snape to look, i.e. to seek perceptual evidence for his assertion that Harry Potter hit his, Malfoy's, friend Goyle in (40-e). And in fact, Snape does look, as is made clear in (40-g). So the support relation between (40-f) and (40-e) belongs to the 'perception' and 'asking to seek' category.[13]

(40) a. Snape pointed a long yellow finger at Malfoy
b. and said, "Explain."
c. "Potter attacked me, sir —"
d. "We attacked each other at the same time!" Harry shouted.
e. "— and he hit Goyle —
f. look —"
g. Snape examined Goyle, whose face now resembled something that would have been at home in a book on poisonous fungi.

In other words, the distinction between different sources of evidence is orthogonal to the distinction between providing and asking to seek. We should also be able to find instances of support relations that provide and ask to seek inferential evidence, thus amounting to six possible types based on these two features.

To summarise, we have distinguished between eight possible types of support relations based on the location of the communicative problem within the cascade of goals, or levels of action, of a speech act (cf. figure 1). We have described five of those types in more detail: (a) support for attention, (b) support for identification, (c) support for understanding, (d) support for uptake, and (e) support for execution. We have made a further distinction between five types of goals according to *what* needs to be achieved, what kind of state or event, giving rise to further types of support relations: (a) support for attention, (b) support for belief, (c) support for desire/intention, (d) support for emotion, and (e) support for action. We have argued that further subdivisions of support relations according to the means the speaker uses to achieve their goal will depend on the type of goal. That is, the means relevant for support for belief and those relevant for support for, let's say, attention will generally not be the same. Finally, we have identified three features that characterise support for belief relations according to

13. Directive speech acts like (40-f) are often accompanied by pointing gestures, but they are distinct from pointing gestures in our taxonomy. Pointing gestures belong to the 'providing' category as they facilitate access to perceptual evidence rather than directly tasking the addressee with an action.

the means used to induce belief: (a) evidence-based vs. pragmatic; (b) providing vs. asking to seek; (c) the source of evidence—perception vs. testimony vs. inference.

Crucially, these are all and only support relations that (should) exist. Of course, one can define more features to further subdivide this space, but for the features discussed so far one could already say that if a relation is a support relation, it must belong to one of the proposed categories at each level, otherwise it is not a support relation. For example, the classical *Evidence* relation (36) is support for uptake, support for belief, evidence-based, providing, and based on testimony. The relation in (40-f)–(40-e) is support for uptake, evidence-based, asking to seek, and based on perception. The relation in (17-a)–(17-c) is support for identification, support for belief, providing, and based on perception. Finally, to give one more example, the relation in (34-a)–(34-c) is support for execution, support for action, and we have not defined further subtypes based on the solution method for support for action relations.

As should have become evident in the meantime, the features this classification is based on are only partly independent. Obviously, attention as a stage in the workflow of a speech act and attention as a type of mental state to be achieved inherently go together. Desire and emotion as types of goals only seem to be relevant for support for uptake relations. However, belief constitutes the goal at several stages—identification, understanding, and uptake. These asymmetries result from our concept of the levels of action in communication. The fact that there are (probably) no support relations that belong both to the support for understanding and support for desire category is a consequence of the nature of understanding—that the result state of understanding is a certain kind of belief, rather than desire. In other words, the asymmetries are exceptions that prove the rule as they are a logical consequence of our top-down, concept-driven approach.

Contrary to tradition, we have not given names (e.g. *Evidence*, *Motivation*, *Enablement*) to all possible combinations of the proposed features, and we would indeed need too many to label all the boxes we have opened. If a short label is needed for ease of reference or for the purposes of corpus annotation, one way to go could be to extend the existing labels to broader categories. For instance, one could agree to refer to all evidence-based support for belief relations as *Evidence*. However, one should keep in mind that this is neither RST nor SDRT *Evidence* any more, and that it covers cases like (40-f)–(40-e) and (17-a)–(17-c) that were not thought of as *Evidence* relations previously. Users of this taxonomy are welcome to come up with relation names as needed.

5. Conclusion and outlook

We started out by expressing our general dissatisfaction with previous definitions of speech-act-level coherence relations which did not specify clearly what it is about speech acts that make some combinations of them coherent while not others. We are now in a position to give a partial answer to that question. In general, coherence between speech acts can be characterised in terms of relations between the goals of those speech acts. There is a limited number of ways in which goals can be coherently related. In this paper, we have investigated one such way—support—where one speech act fails or is anticipated to fail to achieve one of its goals, while the other helps achieve that goal.

Based on theoretically motivated considerations about levels of action in communication and the nature of different types of goals, we have offered some further subdivisions of the broad category of support relations and have shown how some speech-act-level relations known from previous studies, e.g. self-repairs or RST's presentational relations, fit into the proposed categories. However, unlike RST, our radically top-down approach also made it possible to predict what other types of support relations *should* exist, even if they might not be common in written texts—the type of data that predominantly informed early relational approaches to discourse structure. Moreover, for many of the predicted types we were able to find naturally occurring examples, some of which would be really hard to categorise within any previously proposed classifications. This is especially true for sequences that include non-assertive speech acts, which have been barely taken into account by approaches to discourse structure based on coherence relations so far. This is why the proposed taxonomy is particularly relevant for the study of coherence relations in dialogue and will likely turn out useful for the annotation of dialogue corpora.

Another important advantage of the top-down approach is that it defines the limits of what is theoretically possible. Our ultimate goal is to have a comprehensive taxonomy of coherence relations at speech act level, so that if a sequence of speech acts does not fit one of the categories in that taxonomy, then we should be able to say that the sequence should therefore be incoherent. The taxonomy proposed in this paper is not comprehensive. If a sequence does not fit our definition of a support relation, it only means that it is not a support relation, but it could still be coherent at speech act level because it features a speech-act-level relation of a different type. However, it is crucial that the range of other possibilities is also theoretically motivated and very limited. Apart from support relations, we mentioned adjacency pairs and Franke's (1990) retractions as other ways in which speech acts can be related. Yet another possibility, which we did not mention, is when two speech acts (usually of the same type, e.g. two assertions) are designed to achieve distinct subgoals of a bigger communicative goal of the speaker. This case has been studied extensively within the approach to discourse structure based on Questions under Discussion (Roberts, 1996; Büring, 2003). As per Roberts' (1996) original intention, Questions under Discussion are one specific way to operationalise the more general notion of communicative goal. In (41), one could say that the speaker's goal is to make a 'big' assertion in which they inform the hearer about who ate what, but they split it up into two 'small' assertions, one to inform the hearer about what Fred ate, and the other to inform them about what Mary ate.

(41) a. Fred ate the beans.
b. Mary ate the eggplant.

Many of the classical semantic or propositional-level coherence relations, e.g. *Parallel* in (41), *Contrast*, and *Narration*, would fall into this category. But then, we believe that that should be all. All coherent sequences of speech acts should be reducible to these few possibilities and combinations thereof, and if they cannot be characterised in these terms, they should be incoherent. It remains a task for future research to review other types of speech-act-level relations along the same methodological guidelines as those we applied to support relations in this paper.

From the above it should be clear that our taxonomy of support relations, or speech-act-level relations more generally, is not intended to replace existing taxonomies of

semantic relations at propositional level. Propositional-level taxonomies would still be relevant particularly for sequences of assertions. One could attempt to reformulate them in speech-act-theoretic terms, but it would not necessarily bring new insight. The notions of causality, contiguity, and resemblance have turned out useful to capture generalisations about relations between assertions just by looking at the relations between the communicated propositions. Our taxonomy becomes relevant where propositional level taxonomies hit their limits, that is, especially for heterogeneous speech act sequences, interactive sequences of more than one speaker, and sequences that do not go according to plan.

Our approach still needs to be tested in an empirical setting. It is true that applying the proposed relation definitions to corpus annotation would require a lot of reasoning with the speakers' cognitive states, which are accessible to the analyst only to a limited extent. However, in cases where coherence at the speech act level is all we have, i.e. where the more familiar relations at propositional level do not seem to apply, hearers are bound to form at least some hypotheses about the speakers' communicative intentions to be able to perceive the sequence as coherent. If hearers can do it, then analysts should be able to do it too. How reliable the recognition of support relations is, both for hearers and for corpus annotators, is another open question.

References

Nicholas Asher and Alex Lascarides. *Logics of Conversation*, volume 9 of *Studies in Natural Language Processing*. Cambridge University Press, 2003. ISBN 9780521650588. URL https://books.google.de/books?id=VD-8yisFhBwC.

John Langshaw Austin. *How to do things with words*. William James Lectures. Oxford University Press, 1962.

Diane Blakemore. The relevance of reformulations. *Language and Literature*, 2:101–120, 1993. doi: 10.1177/096394709300200202. URL https://doi.org/10.1177/096394709300200202.

Daniel Büring. On d-trees, beans, and b-accents. *Linguistics and Philosophy*, 26(5):511–545, 2003. doi: 10.1023/a:1025887707652. URL https://doi.org/10.1023/a:1025887707652.

Herbert H Clark. Managing problems in speaking. *Speech communication*, 15(3-4):243–250, 1994. doi: https://doi.org/10.1016/0167-6393(94)90075-2. URL https://www.sciencedirect.com/science/article/pii/0167639394900752.

Herbert H. Clark. *Using Language*. Cambridge University Press, 1996.

William K. Clifford. The ethics of belief. *Contemporary Review*, 1877.

Jim Davies and David Matheson. The cognitive importance of testimony. *Principia: An international journal of epistemology*, 16(2):297–318, 2012. doi: 10.5007/1808-1711.2012v16n2p297. URL https://doi.org/10.5007/1808-1711.2012v16n2p297.

Martina Faller. Remarks on evidential hierarchies. In David I Beaver, Luis D Casillas Martínez, Brady Zack Clark, and Stefan Kaufmann, editors, *The construction of meaning*, pages 89–111. CSLI Publications, 2002.

Richard Feldman. The ethics of belief. *Philosophy and Phenomenological Research*, 60(3): 667–695, 2000. doi: 10.2307/2653823. URL https://doi.org/10.2307/2653823.

Anita Fetzer. *22. The structuring of discourse*, pages 685–712. De Gruyter Mouton, Berlin, Boston, 2013. ISBN 9783110214383. doi: doi:10.1515/9783110214383.685. URL https://doi.org/10.1515/9783110214383.685.

Wilhelm Franke. *Elementare Dialogstrukturen: Darstellung, Analyse, Diskussion*, volume 101 of *Germanistische Linguistik*. Niemeyer (De Gruyter), 1990. ISBN 978-3-484-31101-5. doi: 10.1515/9783111631905. URL https://doi.org/10.1515/9783111631905.

Gerd Fritz and Franz Hundsnurscher. Sprechaktsequenzen : Überlegungen zur vorwurf/rechtfertigungs-interaktion. *Der Deutschunterricht*, 27(2):81 – 103, 2009. URL http://www.festschrift-gerd-fritz.de/files/publ_hp/fritz_1975_vorwerfen.PDF.

J. Ginzburg. *The Interactive Stance: Meaning for Conversation*. Oxford linguistics. Oxford University Press, 2015. ISBN 9780198722991. URL https://books.google.de/books?id=vO7ioQEACAAJ.

Jonathan Ginzburg, Raquel Fernández, and David Schlangen. Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics*, 7, 06 2014. doi: 10.3765/sp.7.9. URL https://doi.org/10.3765/sp.7.9.

H. Paul Grice. Meaning. *Philosophical Review*, 66(3):377–388, 1957.

Richard J. Hall and Charles R. Johnson. The epistemic duty to seek more evidence. *American Philosophical Quarterly*, 35:129–139, 1998. URL https://api.semanticscholar.org/CorpusID:169194347.

M.A.K. Halliday and R. Hasan. *Cohesion in English*. English Language Series. Taylor & Francis, 2014. ISBN 9781317869603. URL https://books.google.de/books?id=rAOtAgAAQBAJ.

G. Hindelang. *Einführung in die Sprechakttheorie: Sprechakte, Äusserungsformen, Sprechaktsequenzen*. Germanistische Arbeitshefte. De Gruyter, 2010. ISBN 9783110231472. URL https://books.google.de/books?id=Rp1GblSbotQC.

Laurence R. Horn. *A Natural History of Negation*. University of Chicago Press, 1989.

Joris Hulstijn and Nicolas Maudet. Uptake and joint action. *Cognitive Systems Research*, 7(2):175–191, 2006. ISSN 1389-0417. URL https://doi.org/10.1016/j.cogsys.2005.11.002. Cognition, Joint Action and Collective Intentionality.

Katja Jasinskaja and Elena Karagjosova. Rhetorical relations. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann, editors, *The Companion to Semantics*. Wiley, Oxford, 2020. doi: 10.1002/9781118788516.sem061. URL https://doi.org/10.1002/9781118788516.sem061.

Katja Jasinskaja, Yuting Li, Fahime Same, and David Uerlings. Reference and discourse structure annotation of elicited chat continuations in german. In *Proceedings of The 18th Linguistic Annotation Workshop (LAW-XVIII)*, pages 176–187, 2024. URL https://aclanthology.org/2024.law-1.17.

Jeff Jordan. Pragmatic arguments and belief. *American Philosophical Quarterly*, 33(4): 409–420, 1996.

A. Kehler. *Coherence, Reference, and the Theory of Grammar*. Center for the Study of Language and Information Publication Lecture Notes. Cambridge University Press, 2002. ISBN 9781575862163. URL https://press.uchicago.edu/ucp/books/book/distributed/C/bo3626366.html.

Alistair Knott and Robert Dale. Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, 18:35–62, 1994. doi: 10.1080/01638539409544883. URL https://researchers.mq.edu.au/en/publications/using-linguistic-phenomena-to-motivate-a-set-of-coherence-relatio.

Alex Lascarides and Nicholas Asher. Agreement, disputes and commitments in dialogue. *Journal of Semantics*, 26(2):109–158, 2009. URL https://doi.org/10.1093/jos/ffn013.

Claire Lesage, Nalini Ramlakhan, Ida Toivonen, and Chris Wildman. The reliability of testimony and perception: Connecting epistemology and linguistic evidentiality. In *CogSci*, volume 37, pages 1302–1307, 2015. URL https://www.researchgate.net/publication/275582507_The_reliability_of_testimony_and_perception_connecting_epistemology_and_linguistic_evidentiality.

Willem J.M. Levelt. Monitoring and self-repair in speech. *Cognition*, 14(1):41–104, 1983. ISSN 0010-0277. URL https://api.semanticscholar.org/CorpusID:204981295.

William C. Mann and Sandra Thompson. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3):243–281, 1988. doi: 10.1515/text.1.1988.8.3.243. URL https://doi.org/10.1515/text.1.1988.8.3.243.

William C Mann, CMIM Matthiessen, and Sandra A Thompson. Rhetorical structure theory and text analysis. Technical report, University of Southern California, Information Sciences Institute, 1989. URL https://doi.org/10.1075/pbns.16.04man.

Alan Millar. How visual perception yields reasons for belief. *Philosophical Issues*, 21: 332–351, 2011. doi: 10.1111/j.1533-6077.2011.00207.x. URL https://doi.org/10.1111/j.1533-6077.2011.00207.x.

Karen Niven, Peter Totterdell, and David Holman. A classification of controlled interpersonal affect regulation strategies. *Emotion*, 9(4):498–509, 2009. doi: 10.1037/a0015962. URL https://doi.org/10.1037/a0015962.

Kevin N. Ochsner and James J. Gross. The cognitive control of emotion. *Trends in Cognitive Sciences*, 9(5):242–249, 2005. ISSN 1364-6613. doi: 10.1016/j.tics.2005.03.010. URL https://doi.org/10.1016/j.tics.2005.03.010.

Daniel J O'Keefe. Persuasion. In *The Handbook of Communication Skills. 3d edition*, pages 323–341. Routledge, 2006. ISBN 9780761925392. URL https://books.google.de/books?id=e3V6Zen0UGwC.

Janet Pierrehumbert. *Teun A. van Dijk. Text and context: explorations in the semantics and pragmatics of discourse. (Longman linguistics library, 21.) London: Longman, 1977. Pp. xvii+261*, volume 16. 03 1980. doi: 10.1017/S002222670000640X. URL https://doi.org/10.1515/ling.2005.43.1.207.

Nicolas Porot and Eric Mandelbaum. The science of belief: A progress report. *Wiley Interdisciplinary Reviews: Cognitive Science*, 12, 2021. doi: 10.1002/wcs.1539.

Vasudevi Reddy. Moving others matters. In *Moving ourselves, Moving others*, Consciousness and emotion book series, pages 139–162. Philadelphia, 2012. ISBN 9789027241566. URL https://books.google.de/books?id=gwo5HeeuMbMC.

Craige Roberts. Information structure in discourse: Towards an integrated for-mal theory of pragmatics. *Journal of Heuristics*, 1996. URL https://api.semanticscholar.org/CorpusID:11435099.

Kepa Joseba Rodríguez and David Schlangen. Form, intonation and function of clarification requests in German task oriented spoken dialogues. In *Proceedings of Catalog '04 (The 8th Workshop on the Semantics and Pragmatics of Dialogue, SemDial04)*, pages 101–108, 2004. URL https://clp.ling.uni-potsdam.de/publications/Rodriguez-2004.pdf.

Harvey Sacks and Emanuel A. Schegloff. Opening up closings. *Semiotica*, 8(4):289–327, 1973.

Ted Sanders. Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes*, 24:119–147, 01 1997. doi: 10.1080/01638539709545009. URL https://doi.org/10.1080/01638539709545009.

Ted Sanders, Wilbert Spooren, and Leo Noordman. Toward a taxonomy of coherence relations. *Discourse Processes - DISCOURSE PROCESS*, 15:1–35, 01 1992. doi: 10.1080/01638539209544800. URL https://doi.org/10.1080/01638539209544800.

Ted Sanders, José Sanders, and Eve Sweetser. Causality, cognition and communication: A mental space analysis of subjectivity in causal connectives. In Ted Sanders and Eve Sweetser, editors, *Causal categories in discourse and cognition*, pages 20–59. De Gruyter Mouton, Berlin, New York, 2009. ISBN 9783110224429. doi: doi:10.1515/9783110224429.19. URL https://doi.org/10.1515/9783110224429.19.

Marina Sbisà. Cognition and narrativity in speech act sequences. In Anita Fetzer and C. Meierkord, editors, *Rethinking Sequentiality*, pages 71–97. John Benjamins Publishing Company, Amsterdam, 01 2002. doi: 10.1075/pbns.103.04sbi. URL https://doi.org/10.1075/pbns.103.04sbi.

Julian J. Schlöder and Raquel Fernández. Clarifying intentions in dialogue: A corpus study. In Matthew Purver, Mehrnoosh Sadrzadeh, and Matthew Stone, editors, *Proceedings of the 11th International Conference on Computational Semantics*, pages 46–51, London, UK, 2015. Association for Computational Linguistics. URL https://aclanthology.org/W15-0106.

John R. Searle. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge paperback library. Cambridge University Press, 1983. ISBN 9780521273022. doi: 10.1017/cbo9781139173452. URL https://doi.org/10.1017/cbo9781139173452.

J.R. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cam: Verschiedene Aufl. Cambridge University Press, 1969. ISBN 9780521096263. URL https://books.google.de/books?id=t3_WhfknvF0C.

Eve Sweetser. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge Studies in Linguistics. Cambridge University Press, 1990. doi: 10.1017/CBO9780511620904. URL https://doi.org/10.1017/CBO9780511620904.

Maite Taboada. Rhetorical relations in dialogue. In *Discourse across languages and cultures*, Studies in Language Companion Series, pages 75–97. John Benjamins Publishing Company, 2004. ISBN 9789027295262. URL https://www.sfu.ca/~mtaboada/docs/publications/Taboada_RST_2004.pdf.

Jan Theeuwes. Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, 49(1):83–90, 1991.

Peggy A Thoits. Managing the emotions of others. *Symbolic Interaction*, 19(2):85–109, 1996. URL https://api.semanticscholar.org/CorpusID:144933734.

Rob H. J. van der Lubbe and Albert Postma. Interruption from irrelevant auditory and visual onsets even when attention is in a focused state. *Experimental brain research*, 164(4):464–471, 2005. doi: 10.1007/s00221-005-2267-0. URL https://pubmed.ncbi.nlm.nih.gov/15785951/.

Bonnie Webber, Rashmi Prasad, Alan Lee, and Aravind Joshi. The Penn discourse treebank 3.0 annotation manual, 2019. URL https://catalog.ldc.upenn.edu/docs/LDC2019T05/PDTB3-Annotation-Manual.pdf.

Steven Yantis and John Jonides. Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human perception and performance*, 10(5):601–621, 1984. doi: 10.1037/0096-1523.10.5.601. URL https://doi.org/10.1037/0096-1523.10.5.601.

Steven Yantis and John Jonides. Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human perception and performance*, 16(1):121–134, 1990. doi: 10.1037/0096-1523.16.1.121. URL https://doi.org/10.1037/0096-1523.16.1.121.

Linda Trinkaus Zagzebski. *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*. Oxford University Press, 11 2012. ISBN 9780199936472. doi: 10.1093/acprof:oso/9780199936472.001.0001. URL https://doi.org/10.1093/acprof:oso/9780199936472.001.0001.

Jamil Zaki and W Craig Williams. Interpersonal emotion regulation. *Emotion*, 13(5): 803–810, 2013. doi: 10.1037/a0033839. URL https://doi.org/10.1037/a0033839.