

Rhetorical Relations in Dialogue: A Contrastive Study*

Maite Taboada[†]

Simon Fraser University

1. Rhetorical Relations

The study described in this paper is an analysis of rhetorical relations in a set of task-oriented dialogues, in English and in Spanish. For this purpose, I followed Rhetorical Structure Theory, henceforth RST. Rhetorical relations, under different names and with somewhat different characteristics, have been both theoretical constructs and tools in the work of many discourse analysts. This section introduces the notion of rhetorical relations—which are also referred to as coherence, or discourse, relations—and discusses my choice of RST for the analysis.

As Knott and Sanders (1988, p. 135) point out, coherence relations seem to evolve out of a theory of discourse, and “[t]he starting point for any theory of discourse is the oft-noted observation that what we call ‘a text’ is more than just a collection of random sentences”. “More than just a collection of random sentences” means that texts show some internal coherence, that there are some fixed methods for linking one portion of text to another:

The idea is that a choice amongst a finite set of alternatives must be made when juxtaposing two portions of text, and hence that a finite set of *coherence relations* will be sufficient to enable an analysis of every coherent text. (Knott and Sanders, 1998, p.136)

Whether recognized by the interactants or only by the analyst¹, coherence relations are, then, underlying relations among the propositions in a text. The assumption taken in studies of coherence relations is that those underlying relations can be classified into a limited number of types. Coherence relations reflect pragmatic ties, but also intentionality on the part of the writer. ‘Writer’ is the word here, rather than ‘speaker’, because the application of such relations has focused primarily on written texts. This paper pursues the study of spoken language through coherence relations.

This work is part of a larger project that analyzes generic structure in task-oriented dialogue (Taboada, forthcoming). I considered the dialogues in the study as instances of a *speech genre* (Bakhtin, 1986), noting also that they proceed in a clear sequence of stages or phases. Indeed, according to Martin (1992), genres are purposeful, staged, goal-oriented activities in which speakers engage as members of a culture. The components of this analysis include: surface relations (Theme-Rheme progression), underlying relations (rhetorical relations)

* In C. L. Moder and A. Martinovic-Zic (eds.) *Discourse Across Languages and Cultures*. Amsterdam/Philadelphia: John Benjamins. 2004. 75-97.

[†] Author's address: Department of Linguistics, Simon Fraser University. 8888 University Drive Burnaby, B.C. V5A 1S6, Canada. E-mail: mtaboada@sfu.ca

¹ On the psychological validity (or lack thereof) of coherence relations, see Knott and Sanders (1998) and Redeker (in press).

and cohesion, all three as related to the staging of the dialogues. In addition to the results described here, the RST analysis also contains a study of the presence (or absence) of discourse markers to signal particular relations. For the analysis of underlying relations I chose Rhetorical Structure Theory for different reasons. The most important one is that RST has been extensively tested and applied—although only to written texts. A secondary reason is that it provides definitions of all the relations to be considered, thus increasing the reliability of the analysis.

The next sections describe the results of a cross-linguistic analysis of spoken language, in order to show the explanatory power of rhetorical relations. I only considered the initial set of relations provided by Mann and Thompson in the central description of the theory (1988). The extension, and sometimes reduction, of the number of useful rhetorical relations has been a very much debated question—see, e.g., Hovy and Maier (1992)². My intention is to show how a set of relations which could possibly be enhanced is still powerful for three different purposes:

1. Description of spoken dialogue.
2. Comparison between two languages.
3. Application to the description of a new language, Spanish.

This study presents some significant innovations in the application of rhetorical relations to text analysis. First of all, the mode is spoken language. With some exceptions (e.g., Fawcett and Davies, 1992; Redeker, in press), coherence relations and especially within the RST framework have always been applied to written text.

RST has not yet been effectively related to dialogue. It must be expanded beyond written monologue to dialogue and multilogue in order to encompass a fully representative range of the functions of language. This will perhaps be the most significant modification of its present form. (Mann et al., 1992, p. 68)

Fawcett and Davies (1992) propose RST analyses of conversations that cover intra-turn relations, thus considering a turn as a monologue within a conversation. Daradoumis (1996) extends RST to relations across turns. He follows Berry's (1981) and Martin's (1992) exchange model and provides an extended version, Dialogic RST, with new relations to capture the exchange structure of conversation. Daradoumis' work concentrates also on collaborative conversation, especially tutorial dialogues. My work differs from his in that I did not attempt a modification of the original theory, although I do consider relations across turns.

A second innovation in this work lies with the fact that this is a cross-linguistic study, for which there are few precedents, among them the work of Rösner and Stede (1992). Finally, it is applied to Spanish, which has not been, to my knowledge, analyzed employing RST neither in the written nor in the spoken form.

2. Issues in an RST Analysis of Dialogue

Whereas I will not endeavor to provide a full description of RST, I will, in this section, describe how the adaptation of the theory, from written to spoken text, proceeded. A full account, together with relation definitions and text examples, can be found in: Mann and Thompson (1983; 1987; 1988), Mann et al. (1992), Matthiessen and Thompson (1988), or Thompson and Mann (1987).

² A related question, and criticism of RST, refers to the need of multi-level analysis, in order to include informational and intentional aspects of discourse (Moore and Pollack, 1992 ; Redeker, in press).

An RST analysis of a text implies that the text in question is functionally and hierarchically organized. That dialogue, and especially task-oriented dialogue, features a functional component is an intuitively acceptable consideration. The hierarchical organization seems to be more difficult to discover in an ongoing effort of two participants that have not planned this particular conversation before, are not allowed to produce drafts and final versions, and do not know what the other speaker will say—not specifically, at least, although there is some predictability in a familiar genre or type of conversation.

Considering functional and hierarchical organization, there are two different ways in which the analysis can proceed, according to two different points of view. The first point of view is that of the analyst, where the conversation is presented as a product of the interaction of the two speakers. In this view the conversation is an autonomous, independently-functioning piece of text. The analyst is an observer, maybe even a recipient of this final product. We impose a second layer on the dialogue, which results in a detachment from the context where it happened to the context of the analysis. If this detachment happens at the expense of the original context, how is it useful? Although away from its original context and in a different one, the analysis might provide insights into how the interactants construct a piece of text in a joint manner, taking into consideration what has happened so far in the dialogue and incorporating the other speaker's contribution in order to modify or adjust their own. Mann and colleagues (Mann et al., 1992, p. 41) consider “some dialogues as structured collections of texts”. It is that structural relation that I will try to unveil here: how the dialogue as text is organized and what general characteristics we can find in a corpus of such texts.

The second point of view from which we could look at these dialogues is that of the interactants. In this view, the conversation is a process to which both speakers contribute in their respective turns. Each turn would be an independently created text, a response to the overall context, but a text in its own right. This type of analysis will reveal how and when each turn is autonomous, and how speakers create mini-pieces of discourse with the goal of contributing to a dialogue.

Another consideration is that RST is, in a sense, a theory of influence; it is very well suited to describe what the writer/speaker does to influence the reader/hearer. A look at the inventory of relations will suffice, especially the ones grouped under the presentational heading, “those whose intended effect is to increase some inclination in the reader” (Mann and Thompson, 1988, p. 257)³. The relations can often be described “in terms of the purposes of the writer, the writer's assumptions about the reader, and certain propositional patterns in the subject matter of the text.” (Mann et al., 1992, pp. 44-45).

I will show that in this corpus we can also identify purposes, not of the writer, but of the speaker, who makes decisions in presenting his or her material in a particular way to achieve an effect. The choice and the presentation might not be as elaborate as it is in written discourse, due to time constraints and the impossibility to revise the material. We can, nonetheless, notice options of organization on the part of the speaker. The turn-by-turn analysis considers each utterance as a text in itself that seeks some reaction—an immediate one, unlike in written discourse. This type of analysis ignores the undoubtedly important relations holding between turns, but approximates the original RST analyses in that it approaches text as the product of one mind that projects towards a recipient.

³ Presentational relations are: Motivation, Antithesis, Background, Enablement, Evidence, Justify and Concession.

3. The Study

The corpus used is a total of 60 conversations between two speakers, 30 in English and 30 in Spanish. For each language, the conversations are broken down in ten female-female, ten male-male, and ten female-male dialogues. The dialogues are part of a larger corpus collected by the Interactive Systems Laboratory of Carnegie Mellon University as part of JANUS, a speech-to-speech machine translation project⁴. They were recorded by volunteers, who received a small compensation for their time. The participants were all native speakers of either English or Spanish who lived in Pittsburgh. After recording, the dialogues were transcribed with human and non-human noises and intonation, and divided into semantically independent units, or Semantic Dialogue Units (SDUs). Each SDU is separated by a Semantic End of Segment (SEOS) marker. SDUs are comparable to Chafe's (1980) *idea units*, or to the *turn-constructive units* of Conversation Analysis (Sacks, Schegloff and Jefferson, 1974). The translation of Spanish examples provided throughout the paper is a free translation, rendered one SDU at a time. The translations do not include intonation or non-linguistic sounds (i.e., sounds that are not words). For those, the reader is referred to the original Spanish text.

A summary of the most common symbols in the examples contained in this paper can be found in the Appendix. The conversations are labeled with the name of the speakers that took part in them, followed by an indicator that represents either their order, or the month where they were recorded. The speakers are identified through an M for male or an F for female followed by their initials. Whenever I know the gender of the speaker, I will use the appropriate pronoun to refer to him or her.

Tables 1 and 2 provide a summary of the length of the dialogues, in terms of turns, SDUs and words. In Table 1, the figures represent total counts for each language. In Table 2, the figures provided are all means over the 30 dialogues for each language. We can see that the dialogues tend to contain approximately the same number of turns. However, the Spanish turns are longer in terms of SDUs and words.

Table 1. Counts in the corpus

	English	Spanish
Turns	249	248
SDUs	784	1294
Words	6804	9112

Table 2. Mean lengths of dialogues

	English	Spanish
Mean length - Turns	8.30	8.27
Mean length - SDUs	26.13	43.10
Mean length - Words	225.93	302.43

⁴ My thanks to the Interactive Systems Laboratory of Carnegie Mellon University and its director, Alex Waibel, for permission to access the corpus used in this study. Sondra Ahlen also provided extra information on the data.

As explained in the previous section, two different types of analyses were carried out. In this section I describe the results for each of them in both languages, listing the relations found and providing examples.

3.1. Turn-by-turn Analysis

The turn-by-turn analysis showed a number of repeated patterns in turns sharing the same purpose, elements that were always present because they define the genre of task-oriented dialogue, or because they are characteristic of dialogue. Table 3 displays a list of the relations found in the corpus, for Spanish and English. The percentage column shows the percentage of a particular relation with respect to the total number of relations found in that language.

The results for English and Spanish, in terms of relations and their distributions, are very similar. In both cases, Elaboration shows the highest number of occurrence, followed by Concession and Condition. Lower in the list the relations vary only minimally in percentage distribution. No statistical tests of comparison between languages were performed, given the wide variability of the data, and the presence of counts of zero for some of the relations. Comparisons, then, are tentative and based solely on the percentages with respect to the total for each language.

Table 3. Number and percentages of turn-by-turn analysis

	English	% English	Spanish	% Spanish
Antithesis	1	0.17	1	0.15
Background	13	2.24	16	2.44
Circumstance	5	0.86	9	1.37
Concession	71	12.24	70	10.67
Contrast	9	1.55	8	1.22
Condition	66	11.38	89	13.57
Elaboration	166	28.62	140	21.34
Enablement	10	1.72	4	0.61
Evaluation	8	1.38	11	1.68
Evidence	0	-	4	0.61
Interpretation	1	0.17	11	1.68
Joint	29	5	26	3.96
Justify	34	5.86	22	3.35
Motivation	3	0.52	0	-
Non-volitional Cause	37	6.38	54	8.23
Non-volitional Result	43	7.41	29	4.42
Otherwise	4	0.69	4	0.61
Purpose	10	1.72	26	3.96
Restatement	28	4.83	47	7.16
Sequence	7	1.21	20	3.05
Solutionhood	7	1.21	0	-
Summary	4	0.69	6	0.91
Volitional Cause	7	1.21	11	1.68
Volitional Result	17	2.93	48	7.32
TOTAL	580		656	

The next section provides a detailed analysis of two conversations, where we will also see instances of the intra-turn relations. The remainder of this section will consider some specific problems encountered when analyzing the data. The problems in applying RST to these dialogues had to do with their very nature. In naturally occurring dialogue we find elements pertaining to the specific genre, such as greetings and goodbyes, and also different degrees of spontaneity in the language, such as the presence of self-talk versus the more planned turns in which a speaker's utterance evolves in an easily definable pattern.

The next example shows two different groups of spans within the same utterance. The first part is mainly interactional and not related to the task itself, but to the performance of the task ("well, totally just missed what you said, but if it's a date, oh yeah, yeah"). The second part is related to the task, reflects a reaction to what the previous speaker said, and it consists of three text spans, the first two in an Elaboration relation, and the last one as the nucleus of a Concession relation, as illustrated in Figure 1. (The span numbers displayed in the figure are shown in between square brackets in the example). In this sense, Redeker (in press) remarks that parenthetical expressions that do not constitute background information, especially in spoken language, are not satisfactorily explained in any theory of coherence relations.

(1) [FCAD_FJAB_12]

fjab_12_02: #tap# well {comma} /h#/ /um/ totally just missed /begin_lg/ what you said {comma} /end_lg/ {seos} /h#/ but if it's a date {comma} oh yeah {period} yeah {period} [1] I'm out of town then {period} {seos} [2] I'm out of town from Monday through Wednesday {period} {seos} /h#/ [3] but /um/ {comma} /h#/ /s/ the next week I'm free {comma} /h#/ for the most of the time {period} /h#/ #key_click# /lg/ {seos}

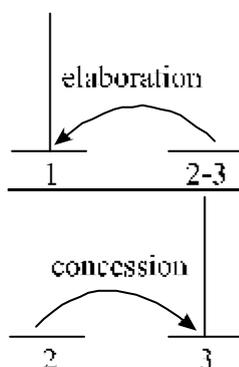


Figure 1. Relations in Example 1

Another problematic case is that of parallel relations between two turns of talk. In Example (2), the two underlined segments hold a relation with each other, but no relation with the rest of the turn itself. In those cases, I decided to divide the turn into two independent groups of relations. The connections will play a role in the analysis of the conversation as a whole, but not in the turn-by-turn analysis.

(2) [FBNT_MJFG_1]

fjab_12_04: /s/ /h#/ well Mrs Daley {comma} /s/ I {comma} happen to /eh/ be in a seminar until six o'clock {comma} on Friday {comma} so I really have a true excuse {period} {seos} /s/ /h#/ but {comma} I like <the {comma}> /um/ the Monday thing {comma} /h#/ two weeks from now {comma} the twenty sixth {comma} {seos} it's a good time for me {period} /h#/ #key_click# {seos}

fcad_12_05: #begin_electronic_hum# /s/ /h#/ that's #tap# Ms Daley to you {comma} {seos} and {comma} /s/ /h#/ #paper_rustle# what time would you like to have it on the twenty sixth {comma} {seos} how 'bout {comma} /s/ /h#/ say {comma} /s/ /h#/ that lunch thing again {comma} #paper_rustle# {seos} you can buy me lunch again at like {comma} twelve to two {quest} {seos} that sounds good {period} /s/ #end_electronic_hum# {seos}

The next aspect—which is not only an aspect of conversation, but also of written texts—pertains to genre-related aspects of the dialogues. The purpose of some stretches of talk seems to be merely that of complying with the characteristics of dialogue in general, or task-oriented dialogue in particular. These particular stretches of talk correspond to the holistic structure of the text “deriving from the genre or variety of text” (Mann et al., 1992, p. 41). Turns in which only greetings or goodbyes are exchanged do not show any coherent organization in relational terms; they merely serve to establish or maintain the communication, and to ascribe the texts to a specific genre, which both speakers know and exploit. Example (3), in Spanish, shows the end of a conversation where speaker FJGC confirms the appointment, and speaker FKAS agrees and signals that the conversation is over.

(3) [FJGC_FKAS_04]

fjgc_04_04: *pause* #paper_touch# /h#/ /um/ el veintiséis no hay problema {period} {seos} /um/ nos podemos reunir de once a una de la tarde {period} {seos} /h#/ está bien {quest} /s/ #key_click# {seos}

‘the twenty-sixth there is no problem {seos} we can meet from eleven to one in the afternoon {seos} is that good’

fkas_04_05: /h#/ sí {seos} está bien {period} {seos} perfecto {period} {seos} hasta luego {period} #key_click# {seos}

‘yes {seos} that's good {seos} perfect {seos} see you later’

I will finish this section on the peculiarities of a turn-by-turn RST analysis for conversation pointing out that some of the problems already described through examples of short turns were also present when analyzing the conversation as a unified, coherent whole. In the analysis of full conversations, also encountered were holistic elements, related to the dialogue as interaction, and thus not considered for the analysis. The same holds for instances of self-talk.

3.2. Conversation-as-a-Whole Analysis

Analyzing the rhetorical relations in a whole conversation entails a certain detachment from the conversation as interaction, as process, towards conversation as finished product. This second type of analysis does not pay so

much attention to the effects achieved on the hearer, but to the cohesiveness of the text as an entity, leaving aside the fact that there was a target hearer or reader for these texts.

In such a situation the procedure was to scan the conversation for the main purpose of the interaction, and from that macro-relation to proceed onto the analysis of the smaller segments. The span or spans that represented the main point in the interaction were usually the ones that contained the proposal for a meeting. Once the problem was expressed, the rest of the conversation dealt with its solution, that is, with finding an appropriate time for holding a meeting. This main point is the Comprehensive Locus of Effect, the ‘‘portion of the text that represents the essence of the text as a whole’’ (Mann et al., 1992, p. 61).

Obviously, at the lower levels of analysis, the relations holding were the same as those found for the turn-by-turn analysis. At the higher level, however, there were new relations between the turns uttered by the different speakers. That is to say, a speaker might create a relation that includes one or more spans previously uttered by the other speaker. Segmentation was not always based on turns (i.e., a segment of text did not always start and end with a turn), although the turn boundaries frequently showed a natural breaking point. The decisions taken in the analysis of turns and the analysis of the conversation as a whole influenced each other, because decisions previously taken—for instance, for the internal structure of a turn—biased the analyst’s decision when the conversation was considered as a whole. In order to avoid, or at least balance, this influence, out of the 30 conversations, 15 were first analyzed as whole conversations, and then broken down in turns. Then the other 15 were analyzed in the reverse order.

I present here one example of analysis of two full conversations, one for each language. Example (4) shows a complete sample English conversation, and its RST analysis is represented in Figure 2. The spans in the figures are represented in between square brackets in the conversations.

(4) [MAGH_MTNZ_11]

magh_11_01: [1] #key_click# /h#/ /h#/ okay {comma} would you have {comma} two hours {comma} on Thursday afternoon {comma} the third {comma} or Friday {comma} the fourth {period} /ls/ #key_click# {seos}

mtnz_11_02: [2] /h#/ /ls/ /h#/ no {comma} [3] that’s not good {comma} {seos} [4] <I have a {comma} {seos}> I’m actually on vacation {comma} Wednesday {comma} Thursday {comma} and Friday {period} {seos} /sniff/ [5] how about /uh/ {comma} this Tuesday {period} {seos} tomorrow {period} #key_click# #click# /h#/ /h#/ {seos}

magh_11_03: [6] /h#/ no {comma} {seos} [7] /um/ that sounds like this week {comma} doesn’t work {period} {seos} [8] /h#/ next week {comma} the earliest I could do it {comma} would be again {comma} Thursday afternoon {comma} or {comma} anytime Friday {period} #key_click# #noise# {seos}

mtnz_11_04: [9] /sniff/ okay {period} {seos} [10] yeah I can’t do it on {comma} Thursday {comma} {seos} [11] but Friday {comma} in the afternoon would be good {comma} {seos} [12] how about /uh/ {comma} one to three {quest} #key_click# #noise# {seos}

magh_11_05: [13] /h#/ /ls/ that’s Friday {comma} June eleventh {comma} one {comma} to three {quest} {seos} [14] <that’s [f(ine)] {seos}> [15] okay {comma} [16] that’ll work {period} #paper_rustle# #key_click# #paper_rustle# {seos}

mntz_11_06: [17] #key_click# great {period} {seos} [18] let's meet then then {period} {seos}
 [19] #key_click# /ls/ /h# / #paper_rustle# see you {period} #key_click# /h# /h# / {seos}

magh_11_07: [20] #key_click# okay {comma} [21] I guess we're done {period} #click#
 #key_click# #headset# {seos}

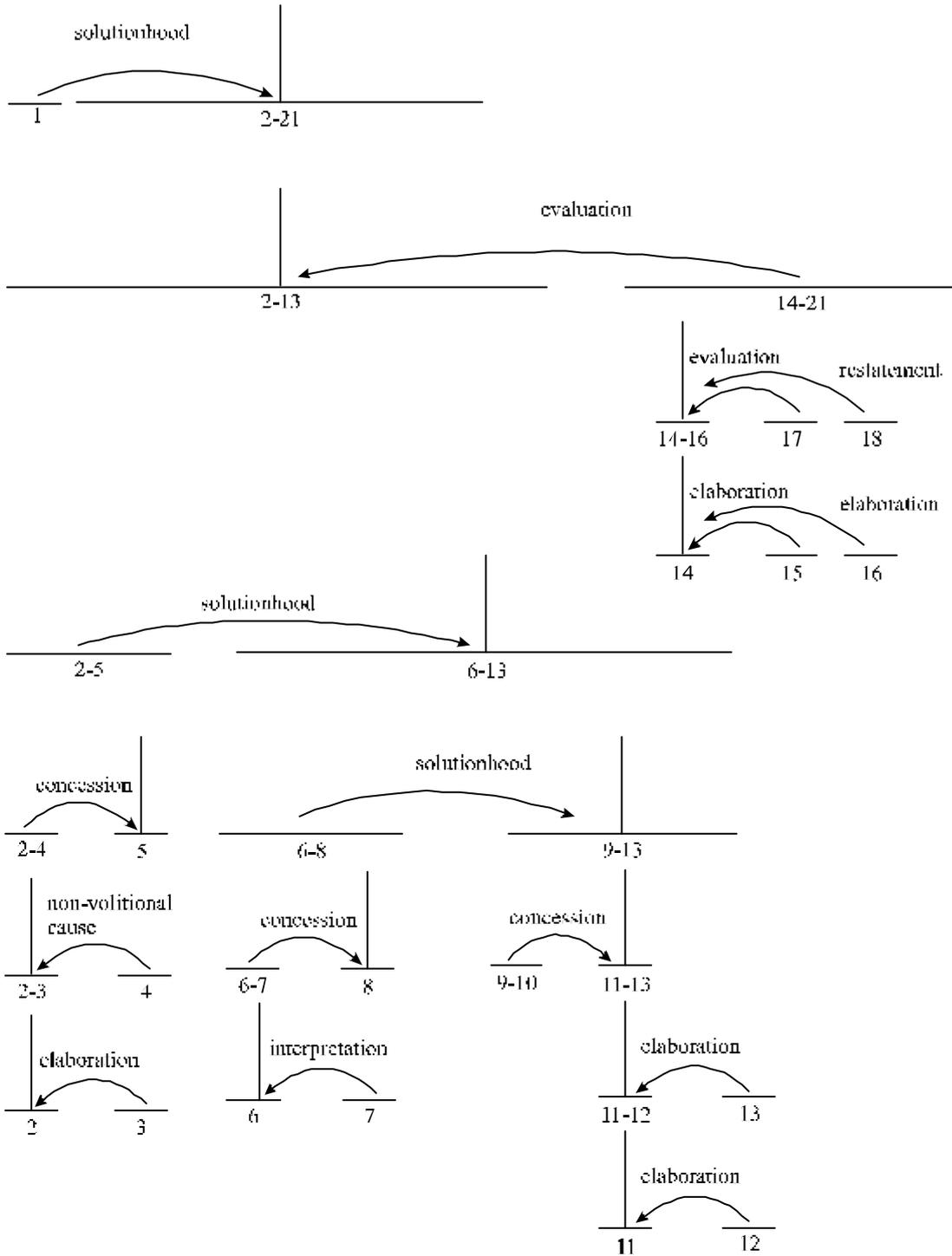


Figure 2. Rhetorical relations in Example 4

The Comprehensive Locus of Effect in this dialogue is a relation of Solutionhood: speaker MAGH asks from speaker MTNZ whether he has free time on a particular date. The problem-posing span is the satellite to the rest of the spans, which provide the solution. The next higher-order division takes place at the moment where speaker MAGH evaluates the current proposal, repeatedly, in spans 14-16: “that’s [f(ine)], okay, that’ll work”. The conversation is divided in two clear sequences at that point, which I will describe now.

The first stage of the conversation, spans 2-13, is organized in another Solutionhood relation, the satellite being a question by MTNZ on a meeting date, and the nucleus the reply to that question. The question itself is composed of a Concession relation with an embedded Cause, which could be paraphrased as “although I can’t on that date because of x reason, how about this other date?” There is also an Elaboration on the negative: “no, that’s not good for me” which softens the bluntness of a single “no”.

Spans 6-13, the nucleus of the reply in the Solutionhood relation, are again broken down into a Solutionhood relation. The satellite is composed of spans 6-8 with a similar structure to the one described above, a Concession and an Interpretation: the “no” of speaker MAGH means that the week under consideration is out of the question and, despite that unavailability (Concession), he is willing to propose another date. Speaker MTNZ uses the same structuring in his reply: a Concession plus a proposal, this time the definitive one. Speaker MAGH elaborates on that proposal in span 13, which is an example of a possible multiple analysis. It could be a Restatement of what has been previously said, or an Elaboration on the date. In this situation, it was decided to assign it to an Elaboration relation because of the addition of new information to the previous speaker’s utterance. We can see how spans 11-12 mention Friday one to three, and speaker MAGH repeats some of that information, but also elaborates on it, adding June the eleventh to the information describing the date.

So far we have seen the description of the first stage of the conversation. The interaction changes its character once the date has been settled. From that point on—spans 14 to 21—there are no more Concession or Solutionhood relations, but Evaluations, Restatements, and Elaborations. Spans 19 to 21 were excluded of the RST analysis because they were considered to be purely interactional, a result of the specific holistic and generic structure of this type of discourse, and then showing no internal RST structure.

In this second part, spans 17 and 18 are the final ones, constituting an Evaluation (“great”) and a Restatement (“let’s meet then”). The previous three spans, 14-16, are simply Elaborations on the Evaluation relation that this second part of the conversation displays with relation to the problem-solving part.

A comparison of the above analysis with Spanish is provided in Example (5), a dialogue whose corresponding RST analysis is shown in Figure 3, and discussed below.

(5) [FSMA_MENF_05]

fsma_05_01: [1] /h#/ hola Edu {period} {seos} [2] mirá {comma} {seos} [3] /eh/ qué te parece si arreglamos /h#/ para /s/ la semana que viene {comma} /eh/ bueno <vernos> /eh/ reunirnos un par de horas {period} {seos} [4] /h#/ /eh/ yo tengo libre el /mm/ lunes veintiséis {comma} después de las doce del mediodía {period} {seos} #paper_rustle# /h#/ /s/ [5] qué te parece {quest} {seos}

menf_05_02: [6] /ls/ /h#/ no {period} {seos} [7] mirá {comma} {seos} [8] yo el lunes no puedo {period} {seos} [9] /eh/ el martes tampoco {comma} {seos} [10] /ls/ /h#/ el miércoles puedo a la tarde {comma} [11] /h#/ y el jueves a la mañana {period} {seos} [12] porqué no me decís qué te parece <en> /h#/ entonces miércoles a la tarde o jueves a la mañana {period} {seos}

*fsma_05_03: [13] /ls/ /h#/ mirá {period} {seos} [14] imposible porque el /ls/ /eh/ miércoles yo tengo de dos a cuatro y media una reunión {comma} {seos} [15] y el jueves a la mañana tengo una conferencia de nueve a doce {period} {seos} /h#/ [16] así que tendría <que> /crky/ que ser recién {comma} /h#/ el miércoles de la semana de [yain (ya_en)] agosto {comma} {seos} el miércoles cuatro {period} {seos} [17] porque yo voy a estar afuera el treinta {comma} /h#/ hasta el tres {period} /h#/ {seos} [18] <y {seos}> bueno {period} {seos} [19] recién sería el miércoles cuatro {comma} /h#/ después de la una del mediodía {comma} o a lo mejor antes de las once de la mañana {period} *pause* {seos}*

menf_05_04: [20] /ls/ /h#/ /eh/ mirá {period} {seos} [21] el miércoles cuatro {comma} por la tarde {comma} yo puedo {period} {seos} [22] así que {comma} si te parece {comma} [23] quedamos el miércoles cuatro {comma} de dos a cuatro de la tarde {period} /h#/ {seos}

fsma_05_05: [24] /ls/ /h#/ bueno {comma} {seos} [25] dále {period} {seos} [26] entonces nos vemos el miércoles {comma} /h#/ /eh/ /mm/ entre las dos y las cuatro {period} {seos} [27] [tá (está)] bien {period} {seos} [28] /h#/ /ls/ chau {period} {seos}

TRANSLATION:

fsma_05_01: [1] hello Edu {seos} [2] look {seos} [3] what do you think if we settle for next week well to <see each other> meet for a couple of hours {seos} [4] I'm free on Monday the twenty-sixth after twelve noon {seos} [5] what do you think {seos}

menf_05_02: [6] no {seos} [7] look {seos} [8] I can't on Monday {seos} [9] on Tuesday either {seos} [10] on Wednesday I can in the afternoon [11] and on Thursday in the morning {seos} [12] why don't you tell me what you think <on> then Wednesday in the afternoon or Thursday in the morning {seos}

fsma_05_03: [13] look {seos} [14] impossible because on Wednesday I have from two to four thirty a meeting {seos} [15] and on Thursday in the morning I have a talk from nine to twelve {seos} [16] so it would have <to> to be then on Wednesday in the week of already August {seos} Wednesday the fourth {seos} [17] because I'm going to be out of town on the thirtieth until the third {seos} [18] <and {seos}> well {seos} [19] then it would be Wednesday the fourth after one in the afternoon or maybe before eleven in the morning {seos}

menf_05_04: [20] look {seos} [21] on Wednesday the fourth in the afternoon I can {seos} [22] so if you want [23] we can meet on Wednesday the fourth from two to four in the afternoon {seos}

fsma_05_05: [24] good {seos} [25] okay {seos} [26] then we meet on Wednesday between two and four {seos} [27] that's good {seos} [28] bye {seos}

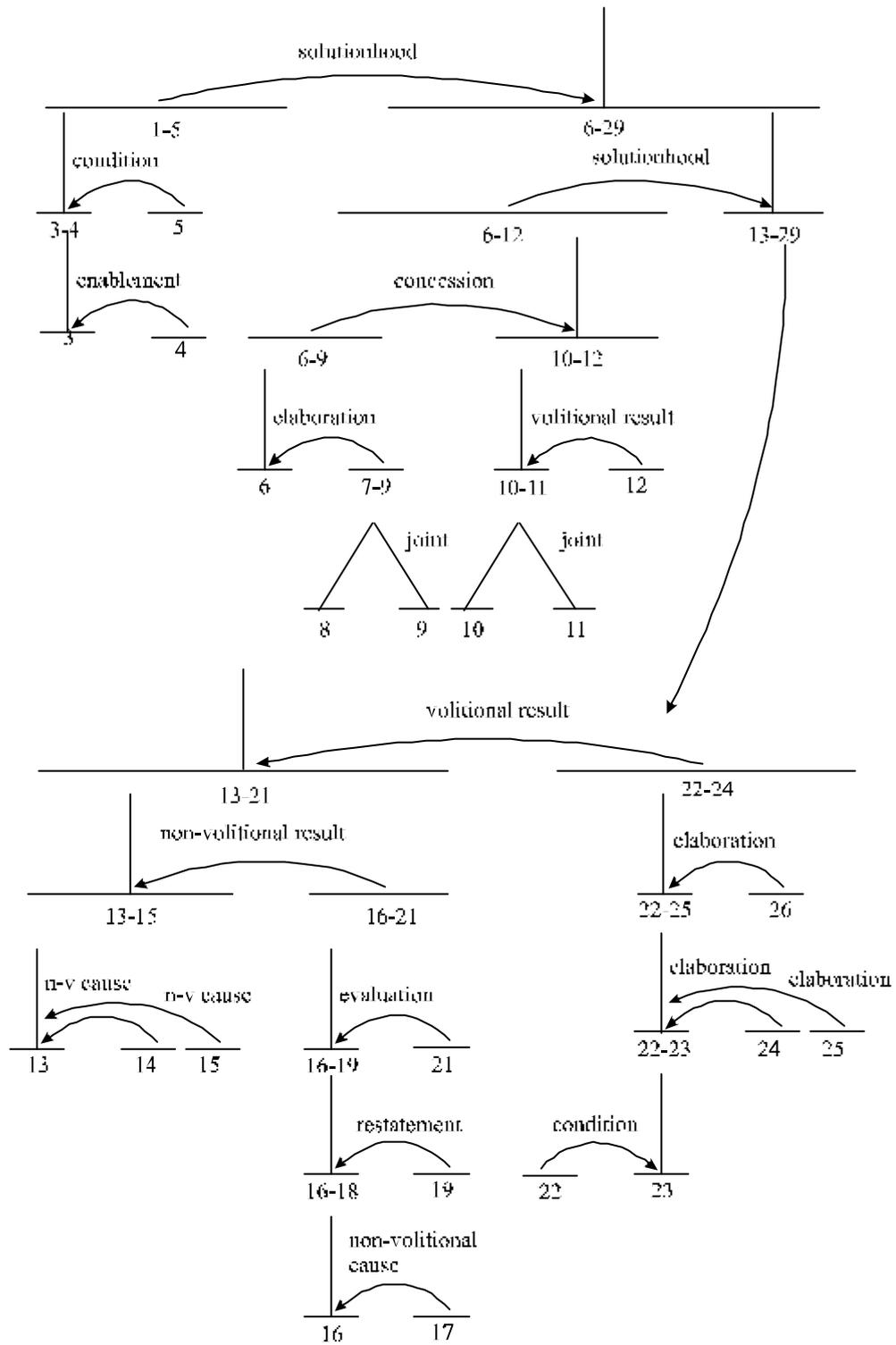


Figure 3. Rhetorical relations in Example 5

The Spanish exemplified dialogue shares many of the structural characteristics we already observed in the English one. It is also built around a Solutionhood relation, with spans 1 to 5 as the satellite and the rest of the dialogue for nucleus. The problem statement includes a Condition relation: span number 5 is the satellite which literally means “what do you think?”, and which I interpreted to be a condition on the proposal presented by the current speaker. A paraphrase could be “I’d like to meet on this date, if you want to/are available”. Span 4 represents an Enablement, the available date to meet, for the proposal of meeting presented in span 3. Spans 1 and 2 were not considered for the RST analysis, because they are merely interactional.

Our next focus of attention is the main body of the Solutionhood relation, the nucleus in spans 6-29. These are again divided into two main parts which stand in a Solutionhood relation to each other. Spans 6 to 12 are the satellite, presenting another question on when to meet, and spans 13 to 29 constitute the answer to that question. I will describe the relations within those two groups of spans separately.

The satellite of the Solutionhood relation is one utterance of speaker MENF, where he presents the constraints on his calendar by means of a Concession relation. He is being collaborative and, as a consequence, although he expresses his unavailability on the date that speaker FSMA proposed, he contributes some other options. The satellite of the Concession relation, spans 6 to 9, consists of an Elaboration on the “no” in spans 8 and 9 which is presented as a joint relation, two nuclei expressing the two dates when he is not available. Span 7 is not considered for the RST analysis. The nucleus of this Concession is presented as a Volitional Result at the top level with another embedded Joint relation. This time the two Joint nuclei present the two free slots of the speaker. As a result of this availability, the speaker would like to know what his interlocutor thinks of those dates. The Volitional Result is marked by *entonces* (‘then’).

Now we turn to explore the structure of the nucleus in the second Solutionhood relation, spans 13 to 29. The first relation we encounter is one of Volitional Result, used again to express the idea that the speaker would like to meet, once the calendars have been described. The nucleus in this relation includes spans 13 to 21 and extends across speakers’ turns, starting at FSMA’s turn (the third turn in the conversation) and continuing onto the fourth turn, that of speaker MENF. This stretch starts with the presentation of a problem: speaker FSMA cannot meet on either of the days proposed. There are two Non-volitional Causes that build the background for the unavailability. The Non-volitional Result of this is that they will have to start considering a week in August for their meeting. The satellite of this Volitional Result includes an Evaluation (satellite in span 21), a Restatement (satellite in span 19), and a Non-volitional Cause (satellite in span 17).

The satellite in the Volitional Result relation—spans 22 to 24—is the acceptance and restatement of the proposed date. As we saw in the English conversation, this last part of the interaction displays a completely different structure, where no Concessions or relations of Cause are present, only Elaborations on and Evaluations of the date proposed.

Table 4 shows the number and percentages found in the analysis of the conversations as whole texts. The most frequent relation is still Elaboration, but this time two other relations are salient in their increase: Evaluation and Solutionhood. There is also a slight increase in the Restatement relation. I proceed now to explain their occurrence and to provide some examples.

Table 4. Number and percentages of the analysis of whole conversations

	English	% English	Spanish	% Spanish
Antithesis	1	0.13	1	0.12
Background	14	1.81	18	2.08
Circumstance	4	0.52	7	0.81
Concession	71	9.16	70	8.09
Contrast	7	0.90	8	0.92
Condition	67	8.64	89	10.29
Elaboration	175	22.58	140	16.18
Enablement	11	1.42	4	0.46
Evaluation	82	10.58	67	7.75
Evidence	0	-	4	0.46
Interpretation	1	0.13	11	1.27
Joint	29	3.74	26	3.01
Justify	34	4.39	22	2.54
Motivation	3	0.39	0	-
Non-volitional Cause	40	5.16	54	6.24
Non-volitional Result	45	5.81	30	3.47
Otherwise	4	0.52	4	0.46
Purpose	10	1.29	26	3.01
Restatement	24	3.10	71	8.21
Sequence	7	0.90	20	2.31
Solutionhood	121	15.61	126	14.57
Summary	4	0.52	7	0.81
Volitional Cause	4	0.52	11	1.27
Volitional Result	17	2.19	49	5.66
TOTAL	775		865	

In the Evaluation relation the effect achieved is that the reader/hearer recognizes that the situation presented in the satellite assesses the situation presented in the nucleus and recognizes the value that it assigns. In the corpus this is mostly an inter-turn relation because the evaluations refer to what the other speaker said previously, and rarely to what the speaker is currently saying. This relation is never signaled by a discourse marker in the corpus.

Example (6) below shows the Evaluation relation in the second turn, that of speaker FACR, responding and evaluating the proposal put forth by speaker FACR.

(6) [FACR_FJYK_AU]

fjyk_au_04: #key_click# /ls/ /h#/ #paper_rustle# /ls/ okay nine thirty in the morning {comma} /h#/ /ls/ /h#/ until {comma} eleven thirty {period} {seos} #paper_rustle# /h#/ /ls/ /h#/ and then perhaps we can #paper_rustle# <grab some> #noise# /um/ get some lunch #click# afterwards {period} {seos} #click# how 'bout {comma} you come over to my office {period} and then {comma} /ls/ /h#/ #noise# perhaps #noise# we can go over to the #begin_noise# conference room that's #end_noise# right next to my office {period} #paper_rustle# #key_click# #noise# #key_click# /h#/ {seos}

facr_au_05: #key_click# /ls/ /h#/ that would be fine on August eighth at #paper_rustle# nine thirty
 {comma} {seos} /h#/ /um/ #paper_rustle# /ls/ I would love to get some lunch as well afterwards
 {comma} /noise/ /noise/ #click# #key_click# #microphone# {seos}

The next predominantly inter-turn relation is Solutionhood. Mann and Thompson (1988, p. 272) define Solutionhood as a situation, expressed in the nucleus, that presents a solution to the problem stated in the satellite. The terms problem and solution are broad, and they cover questions, requests (including requests for information), expressions of needs, and conditions that carry negative values. I have considered the question-answer pairs to be in a Solutionhood relation. Question and answer here are understood, in most cases, as more than one clause, not in the strict sense of two turns containing one clause or sentence each. A problem might be expressed in a long stretch of talk with embedded relations within it, and the same applies for the solution. As with Evaluation, Solutionhood is not signaled through discourse markers, although the most clear marking is the presence of an interrogative clause in the satellite.

The last relation that I will examine is Restatement. Although there are intra-turn Restatements, some of them refer to something mentioned in the previous turn. In English we find only four of these inter-turn Restatements. Unlike in English, in Spanish the number of inter-turn Restatements is very high in comparison to the intra-turn ones. If in English the number went from 24 to 28, in Spanish the intra-turn relations number 47; with the addition of the inter-turn ones, they climb to 71. The Spanish speakers repeat much more often the previous utterances, sometimes even repeating two or three times the same information at the end of a conversation, in a closure sequence of sorts. These repetitions are also marked: 21 out of the 71 relations are marked with *entonces* ('then'). This is indeed one of the most significant differences found in the analysis of the English and Spanish data. In Example (7), we can see how FKAS's turn is a Restatement of the previous speaker's turn.

(7) [FYMM_FKAS_02]

*fymm_02_06: /h#/ {begin_english} /s|okay {end_english} {period} {seos} está bien el martes treinta {period} {seos} a partir de las tres de la tarde {comma} estoy yo desocupada también {period} *pause* #key_click# /h#/ {seos}*

'okay {seos} it's fine Tuesday the thirtieth {seos} from three on I am free too'

*fkas_02_07: /h#/ entonces <nos> nos reunimos a las tres {comma} hasta las cinco de la tarde {comma} martes {period} {seos} /h#/ hasta luego {period} #key_click# *pause* {seos}*

'then we meet at three until five in the afternoon Tuesday {seos} see you'

There are some other instances of inter-turn relations, such as Volitional or Non-volitional Result, but above I have discussed the most relevant ones. We now turn to the occurrence of rhetorical relations in the different stages of the conversations.

4. Rhetorical Relations and Staging

The dialogues included in this study, as instances of a particular genre, proceed in a series of clearly definable stages. I will briefly describe how the stages in the dialogues are represented by, and usually contain, specific relations. It is beyond this paper to discuss genre theory, and how speakers and hearers may come to know that such a conversation should evolve in a *scripted* manner, but the conversations, invariably, evolved in an easily predictable sequence of steps.

The first level of organization is that of a telephone conversation (Schegloff and Sacks, 1973)⁵, also found in other instances of task-oriented dialogue (Maier, 1996). There are, in the dialogues in the corpus, three clear stages: *initialization*, *task-performance* and *closing*. The first stage varies in length and content. It goes from a simple “hello”, or just a throat-clearing noise, to a lengthy exchange involving questions about the other person’s health, work, family, etc. I considered an initialization, or opening, to be any segment of talk from the beginning of the conversation until the need to establish a meeting is stated by one of the speakers. Here there is a remarkable language difference: the openings of Spanish conversations are, on average, 1.2 SDUs (Semantic Dialogue Units) long, whereas the openings of the English conversations are 0.6 SDUs long. For an example, let us look at (8). It is a typical English example, where the first speaker starts with a simple “oh”. Other words, or sounds, include “okay”, “let’s see”, and “um”.

(8) [FBNT_MJFG_01]

fbnt_1_01: #begin_electronic_hum# #click# /s/ /oh/ {comma} would #microphone# you like to meet {comma} #noise# /uh/ {comma} /h#/ #key_click# for a two hour appointment {comma} #click# #squeak# *pause* on {comma} *pause* Monday {quest} at {comma} five #microphone# PM {comma} #key_click# #end_electronic_hum# {seos}

The next example is a typical Spanish initialization. Although not very long—the second speaker immediately asks about the reason of the “call”—the speakers greet each other, use their names, and ask about each other’s well-being.

(9) [MBRP_MFJR_05]

mbrp_05_01: /s/ /h#/ {begin_english} /s/okay /s/Freddy {end_english} {period} {seos} /h#/ cómo estás {quest} /h#/ {seos}

‘okay Freddy {seos} how are you’

mjfr_05_02: bien {begin_english} /s/Brent {end_english} {comma} {seos} cómo estás tú {quest} {seos} para qué me has llamado {quest} {seos}

‘Brent {seos} how are you {seos} why did you call me’

Most of the initializations do not even constitute a full span of text, making it difficult to relate it to any other span. When there are actual relations, those are of Solutionhood, as in Example (9) above, or of Elaboration—a sequence of expressions such as “oh”, “okay”, “well”, “let’s see”.

⁵ The speakers do not face each other, in order to avoid non-verbal communication.

Now we turn to the closing stage, where the speakers usually confirm the date agreed upon, say goodbye, and close the conversation. Closings were considered to begin at the point where one of the speakers' proposal had been accepted by the other speaker. Again, in this stage, the English and the Spanish conversations differ in length and type of relations. Whereas the English are usually short (3.4 SDUs on average), the Spanish closings are longer (5.3 SDUs) and contain a higher number of Restatement and Summary relations, often marked by *entonces*. Example (9) is a typical Spanish closing. The actual closing begins with the *entonces* clause in FSNM's turn.

(9) [FFCS_FSNM_01]

*ffcs_01_09: *pause* /h#/ perfecto {period} {seos} porque no nos reunimos <a la una> /ah/ de una a las tres {period} {seos} qué tal {quest} *pause* #key_click# /s/ /h#/ {seos}*

'perfect {seos} why don't we meet at one /ah/ from one to three {seos} how's that'

*fsnm_01_10: /s/ *pause* será muy bueno esa hora {comma} {seos} entonces te veré allí {period} #noise# #key_click# /h#/ {seos}*

'it'll be good at that time {seos} then I'll see you there'

*ffcs_01_11: /s/ bueno {period} {seos} está bien {period} {seos} te veo el martes a la una {period} {seos} hasta luego {period} *pause* #key_click# *pause* {seos}*

'okay {seos} that's good {seos} I'll see you Tuesday at one {seos} see you'

The task-performance stage itself is composed of different substages. The usual flow of the conversation is represented in Figure 3. There are some variations on this general scheme, but it applies to most conversations, equally in English and in Spanish.

An exemplification can be found in (4) in Section 3.2. The typical realization of the flow in (4) is the sequence of Non-volitional Causes or Results providing reasons for the rejection of the proposal. In the sequence Reason-Rejection, the typical relation is one of Non-volitional Result ("I have another meeting at that time, so I can't meet"). In the sequence Rejection-Reason, the relation is usually one of Non-volitional Cause ("I can't meet on that date, because I already have another meeting"). Non-volitional relations are often used in this context, in order to avoid the loss of face by the other speaker. On the other hand, when an Acceptance takes place, the relations are most frequently Volitional Cause and Result, depending on where the optional Reason is placed.

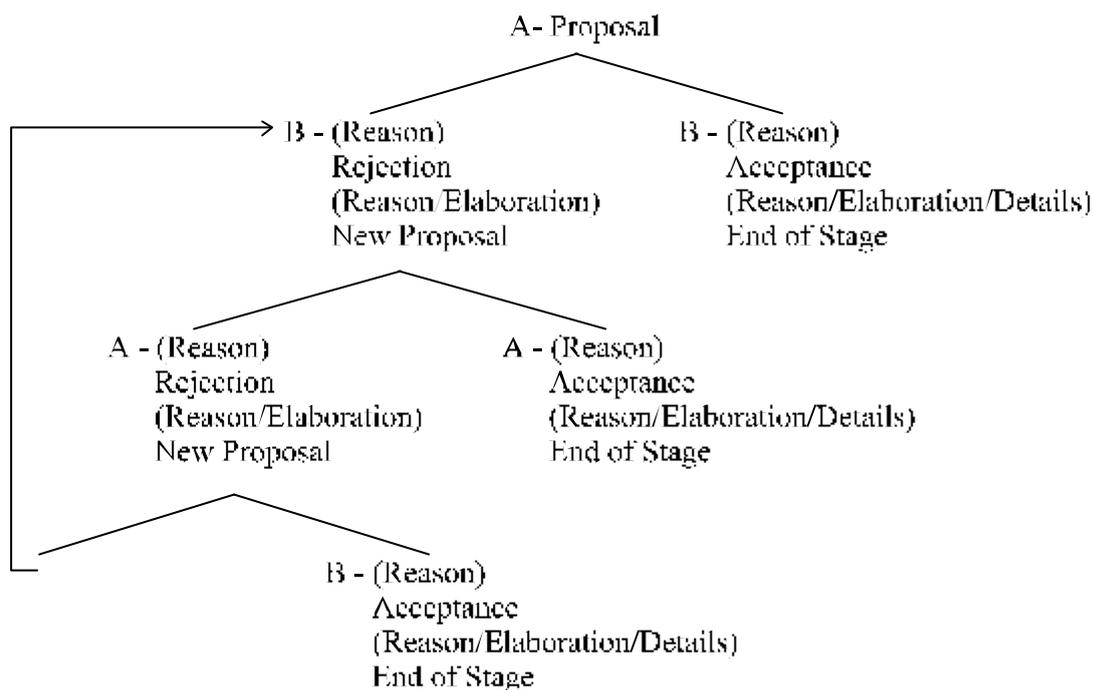


Figure 3. Flow of conversation in the task-performance stage

The other two elements are New Proposal, usually a Result (Volitional or Non-volitional) of the speaker's agenda, followed by a Condition ("we can meet on this other date, if that's okay with you"). Finally, the Details are commonly expressed through Elaboration relations.

I have thus far concluded that both languages share similar characteristics in the structuring of the conversations in three stages. With the exception of longer initializations in Spanish, the length and structure of these stages is also very similar. Furthermore, the types of rhetorical relations used by speakers are similar in similar environments. For instance, the rejection of a proposal will very likely be expressed as a Non-volitional Cause or Result in English and in Spanish. I would like now to raise an issue of generality. One would be tempted to say at this point that the context of situation and the characteristics of the task drive the realization of rhetorical relations to similar extents in the two languages. In other words, given a particular task that involves being cooperative, speakers of English and Spanish avail themselves of similar rhetorical resources. It is my belief that this is true to a certain extent. However, and although these are two different languages, they do not constitute two different contexts of culture. The conversations in the corpus were recorded in the United States, and we need to ponder how much that might have influenced the structuring of the Spanish dialogues. In summary, the task being carried out and the need to be cooperative and polite are certainly all driving forces in the structuring of the rhetorical relations in both languages. The context of culture might be another factor, which cannot be teased apart in this particular study.

5. Concluding Remarks

I have presented the results of a cross-linguistic analysis of spoken language, where I have closely examined the underlying relations between propositions in the discourse of two speakers who are trying to complete a task.

Results show, first of all, that spoken language is suitable for a rhetorical structure analysis, and that patterns in the staging of the dialogues correlate with the type of rhetorical relations present in a particular stage. The cross-linguistic analysis showed no significant difference between the English and the Spanish data in terms of staging or type of relations. The only difference was found in the length of the initialization and closing stages of the dialogues, which tend to be longer in terms of Semantic Dialogue Units (SDUs) in Spanish.

I would like to conclude now with a remark related to the choice of rhetorical relations by speakers. I have already determined how the stage in the dialogue, and the subject matter (e.g., accepting or rejecting a date), in relation to politeness constraints, have an influence on the choice of rhetorical relation. There seems to be another factor at play, namely the preference of any given speaker for a particular relation. If a speaker uses, for instance, a Non-volitional Cause the first time he or she rejects the date, they are very likely to use the same relation to reject the next date. Moreover, this is true of less common relations. One Spanish speaker had a preference for Justify relations, which he used three times in the conversation, the median use for all conversations being 1 and the mean 0.73. The phenomenon occurs across speakers as well. The fact that one speaker used a particular relation results in a higher likelihood of the other speaker's use of that same relation. In a conversation from the English corpus, FHKR_FDMMR_JA, the speakers use Justify four times, three times by speaker FHKR and once by speaker FDMR, after she has heard Justify used twice by the first speaker. The median use for Justify in English is 1.5, and the mean is 1.13. These speakers are, then, using the Justify relation much more often than average, and the only reason that seems to determine its frequency is the presence of a previous Justify relation. Weiner and Labov obtained comparable results in a study of agentless passives (1983). Given a number of factors that influenced the presence of a passive in discourse, the establishment of parallelism with a previous construction in which passive had been used was the highest determining factor. Of course, we should question why *any* relation is used the first time. This paper has discussed the influence of subject matter, politeness constraints and dialogue stages on the choice of rhetorical relations.

REFERENCES

- Bakhtin, Mikhail. 1986. *Speech Genres and Other Late Essays*. Austin: University of Texas Press.
- Berry, Margaret. 1981. "Towards layers of exchange structure for directive exchanges." *Network 2*: 23-32.
- Chafe, Wallace. 1980. "The deployment of consciousness in the production of a narrative." In Wallace Chafe (ed.) *The Pear Stories: Cognitive, Cultural and Narrative Aspects of Narrative Production*. Norwood, NJ: Ablex, 9-50
- Daradoumis, Thanasis. 1996. "Towards a Representation of the Rhetorical Structure of Interrupted Exchanges." In Giovanni Adorni and Michael Zock (eds.) *Trends in Natural Language Generation: An Artificial Intelligence Perspective*. Berlin: Springer Verlag, 106-124.
- Fawcett, Robin P. and B. L. Davies. 1992. "Monologue as a turn in dialogue: Towards an integration of exchange structure and Rhetorical Structure Theory." In Robert Dale, Eduard Hovy, Dietmar Rösner

- and Oliviero Stock (eds.) *Aspects of Automated Language Generation*. Springer-Verlag: Berlin, 151-166.
- Hovy, Eduard and Elisabeth Maier. 1992. *Parsimonious of Profligate: How Many and Which Discourse Structure Relations?* Information Sciences Institute Technical Report. ISI/RR-93-373.
- Knott, Alistair and Ted Sanders. 1998. "The classification of coherence relations and their linguistic markers: An exploration of two languages." *Journal of Pragmatics* 30 (2): 135-175.
- Maier, Elisabeth. 1996. "Context Construction as Subtask of Dialogue Processing: The VerbMobil Case." In *Proceedings of the Eleventh Twente Workshop on Language Technology, TWLT 11*.
- Mann, William C. and Sandra A. Thompson. 1983. *Rhetorical Propositions in Discourse*. Information Sciences Institute Technical Report. ISI/RR-83-115.
- Mann, William C. and Sandra A. Thompson. 1987. *Rhetorical Structure Theory: A Theory of Text Organization*. Information Sciences Institute Technical Report. ISI/RS-87-190.
- Mann, William C. and Sandra A. Thompson. 1988. "Rhetorical Structure Theory: Toward a functional theory of text organization." *Text* 8 (3): 243-281.
- Mann, William C., Christian Matthiessen and Sandra A. Thompson. 1992. "Rhetorical Structure Theory and Text Analysis." In William C. Mann and Sandra A. Thompson (eds.), *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text*. Amsterdam and Philadelphia: John Benjamins, 39-78.
- Martin, James. 1992. *English Text: System and Structure*. Amsterdam and Philadelphia: John Benjamins.
- Matthiessen, Christian and Sandra A. Thompson. 1988. "The structure of discourse and 'subordination'." In John Haiman and Sandra A. Thompson (eds.), *Clause Combining in Grammar and Discourse*. Amsterdam and Philadelphia: John Benjamins, 275-329.
- Moore, Joanna and Martha Pollack. 1992. "A problem for RST: The need for multi-level discourse analysis." *Computational Linguistics* 18 (4): 537-544.
- Redeker, Gisela. In press. Coherence and structure in text and discourse. In William Black and Harry Bunt (eds.) *Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics*. Amsterdam: John Benjamins.
- Rösner, Dietmar and Manfred Stede. 1992. "Customizing RST for the Automatic Production of Technical Manuals." In Robert Dale, Eduard Hovy, Dietmar Rösner, and Oliviero Stock (eds.) *Aspects of Automated Language Generation*. Springer-Verlag: Berlin, 199-214.
- Sacks, Harvey, Emmanuel Schegloff and Gail Jefferson. 1974. "A simplest systematics for the organization of turn-taking in conversation." *Language* 50: 696-735.
- Schegloff, Emmanuel and Harvey Sacks. 1973. "Opening up Closings." *Semiotica* 8: 289-327.
- Taboada, Maite. Forthcoming. *Collaborating through Talk: The Interactive Construction of Task-Oriented Dialogue in English and Spanish*. Ph.D. Dissertation. Universidad Complutense de Madrid.
- Thompson, Sandra A and William C. Mann. 1987. "Rhetorical Structure Theory: A Framework for the Analysis of Texts." *IPRA Papers in Pragmatics* 1: 79-105.

Weiner, Judith E. and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19: 29-58.

APPENDIX

Transcription Conventions

The transcripts include a number of conventions introduced by the transcriber. These include human and non-human noises, as explained below.

CATEGORY	BRACKET
human noises	/.../ slashes
non-human noises	#...# hash marks/pound sign
silences	*...* asterisks
mispronunciations	[...] square brackets (around whole word)
	(...) parentheses (supply missing part of word or correct pronunciation of word, only inside square brackets)
transcriber comments	{...} curly braces
accent	... vertical bars/pipes
false starts	<...> angled brackets

In addition, transcriber comments include intonation, marked with one of the following at the end of the corresponding section of speech.

{period}	Falling intonation
{comma}	Slightly rising intonation, continuation of idea, and not a question
{quest}	Marked rising intonation

These comments do not reflect, or are influenced by, sentence structure. The speaker may have the intonation of a statement whether he or she is, in fact, asking a question. He or she may have the intonation of a period after a collection of words that do not, in any way, resemble a grammatically correct or complete sentence.