Spontaneous and non-spontaneous turn-taking[*]

Maite Taboada

Department of Linguistics

Simon Fraser University

8888 University Dr.

Burnaby, B.C. V5A 1S6

Canada

mtaboada@sfu.ca

Abstract

Turn-taking is usually considered to follow a simple set of rules, enacted through a perhaps more complicated system of signals. The most significant aspect of the turn-taking process is that, in most cases, it proceeds in a very smooth fashion. Speakers signal to each other that they wish to either yield or take the turn through syntactic, pragmatic, and prosodic means. In this paper, I explore how the turn-taking process develops in two different sets of Spanish conversations. In the first group of conversations, speakers take turns spontaneously, presumably as they would do in everyday situations. In the second group, turns were mechanically controlled, and communication was one-way. A comparison of the two types of conversation provides insights into the signals used in spontaneous turn-taking.

---

1. Taking turns when talking

Goodwin (1981: 24), reporting on a comparison by Jaffe and Feldstein (1970), proposes that everyday conversation is similar to short-wave radio as to how the turn-taking is performed. The speaker provides an end-of-message signal, after which the hearer holds the channel, bringing about a change in the speaker/hearer roles. In one-way short-wave radio communication, this end-of-message signal is verbalized in a pre-established word, in English usually "over". The difference between the two types of interaction is that, in a normal conversation, speakers avail themselves of other means or mechanisms to provide that end-of-message signal. My purpose in this paper is to explore which exactly are those mechanisms that speakers use in order to signal turn-taking, with a focus on Spanish.

This study is an analysis of a corpus that contains conversations between dyads of native speakers of Spanish. The conversations are task-oriented—participants have to agree on scheduling an appointment—and were recorded using two different procedures. In a subset of the conversations, the speakers used a computer to indicate the end of a turn. When they had finished their turn, they press the "Enter" key on a keyboard to yield the floor to their interlocutor. This group of conversations is comparable to conversations via one-way communication channels. In the other subset of the corpus, turns are spontaneous, without the need for mechanical intervention. In neither case did speakers establish visual contact, making it impossible to communicate turn-taking through gaze or other gestures. As a consequence, in the second group of conversations, the speakers must supply other verbal signals in place of the "over" signal. The signals may range from pauses (filled or unfilled) or discourse markers to questions addressed at the interlocutor.

In this paper I describe the analysis carried out in both types of conversation. The focus of attention is on turn-constructional units (Sacks et al., 1974), and the ending and beginning portions of a turn (the turn-assignment components). Through an examination of the resources available to speakers, I establish a preliminary description of how Spanish speakers signal turn-taking in spontaneous conversation. The corpus used for

this study happens to be a most appropriate tool to test Goodwin's observation. The recording situation is exactly the same for both types of conversations, the only difference being the mediating of the turn-taking through a computer keyboard. Section 3 describes the data collection and other characteristics of the corpus. The next section reviews previous studies of turn-taking mechanisms, and Section 4 examines the characteristics of turn-taking and turn-yielding in the corpus. Section 5 ends the paper with a discussion of the phenomena observed.

2. Studies in turn-taking behaviour

Sacks, Schegloff and Jefferson (1974) observed a number of characteristics in conversation, among them: variable turn order and size; variable distribution of turns; overlapping is common, but brief; and overlapping is promptly repaired (when two parties find themselves speaking at the same time, one of them will stop). Given these characteristics, it is obvious, according to Sacks et al., that turn-allocation techniques are being used. The current speaker may select a different next speaker, or either party may self-select. This paper is concerned with those turn-allocation techniques that are "obviously" present.

Sacks et al. (1974) propose a set of rules that apply at each *transition-relevance place*, that is, at the point where a next turn can be expected. At each transition-relevance place, choices are presented to both speaker and hearer(s) as to who is to utter the next turn-constructional unit. However, Sacks and his colleagues did not detail the signals employed in communicating such choices. Their rules only specify that, at any given transition-relevance place, the turn-so-far might be "so constructed" as to involve, or not to involve, the use of speaker self-selection. In summary, turn-taking consists of a set of rules, which are enacted through the use of signals (Duncan, 1972, 1973).

Here, I refer to a turn, and to turn-taking, as instances of floor-taking with the intention of holding the floor for a certain period of time. Thus, a turn is different from the situation where a speaker produces backchannel signals (Yngve, 1970). Backchannel signals, such as *uh-huh*, *right*, *yeah*, etc., are signals that the channel is still open, and

they indicate at the same time that the listener does not want to take the floor. Duncan (1972) also establishes a distinction between simultaneous turns and simultaneous talking. Instances of the first involve true overlapping, whereas instances of simultaneous talking do not always imply that the current hearer intends to take the turn; they might just be the result of backchannel signals overlapping with the current speaker's turn.

In the rest of this section, I discuss the nature of a turn, and the different possible signals and rules that have been proposed to account for the observed turn-taking behaviour in conversation.

2.1 *What is a turn?*

According to Edelsky (1981), turn definitions can be grouped in two main camps: mechanical and interactional (see also Furo, 2001). The first group treats turns as units of talk in interaction, without taking into account social context. In this group are studies by Jaffe and Feldstein (1970) and Duncan and Fiske (1985), because for them the turn merely consists of talk with an end boundary. Turns are attributed to a single speaker and are defined in terms of the behaviour of other parties in the conversation (a turn ends when somebody else claims the floor). Goffman (1981) says that a turn is the opportunity to hold the floor, not necessarily what is said while holding it.

On the other hand, interactional definitions are concerned with what happens during the interaction, and take into consideration the intention of the turn taker. Edelsky (1981) points out that speakers are more concerned with completing topics than structural units. Therefore, she defines turn as instances of on-record speaking, with the intention of conveying a message. She also differentiates *turn* and *floor*, since it is often difficult to determine who has the floor, such as situations where a turn is constructed collaboratively by more than one speaker. The floor is the activity taking place or the topic being discussed, often done in collaboration. Hayashi (1991) expands on this definition of floor, describing it as a means of orientation to the communication at the higher level of conversation structure.

Selting (1998; 2000) carries out an extensive review of the meaning behind the concept of turn-constructional unit (TCU) as a unit of talk. TCUs were proposed by Sacks et al. (1974) as the basic units of conversation. Each TCU ends in a transition relevance place, that is, the place where the turn may shift to another speaker. Selting characterizes the notion of TCU as holistic and in need of interpretation. Study of TCUs is relevant here, because one TCU may constitute a complete turn. TCUs may be as short as a word or as long as a sentence. Selting discusses the criteria to divide a turn into units, and concludes that it requires examination of both syntactic and prosodic components.

As we will see in the next section, units in conversation are defined by their boundaries: a unit is talk produced up until an end-point, the point where another interlocutor can take the floor. Ford and Thompson (1996) defined units as those characterized by ending in a complex transition relevance place. They added "complex" to Sacks et al.'s (1974) transition relevance places, because turn units were found to be identifiable through the complex interaction of syntactic, prosodic and pragmatic signals. One interesting aspect of their study is that they used backchannels and laughter by the interlocutor as a signal of a possible completion point of the current speaker's turn. Backchannels by the interlocutor are produced at a point where the current speaker could finish their talk, i.e., a (complex) transition relevance place. That is certainly the case: backchannels are produced at a point where an interlocutor *could* take the turn, but with the backchannel the interlocutor signals that they do not want to do it, in addition to signalling their understanding of or agreement with what is being said (Schegloff, 1982; Yngve, 1970). Ford et al. (1996) also present an extensive discussion on the definition of a unit in talk, concluding that TCUs are emergent, rather than pre-defined, and that syntax, prosody and gesture all contribute to defining the basic unit of talk.

In this paper, I define a turn as continuous talk by one speaker, uninterrupted by the other speaker. There may be talk by the other speaker, but that is often in the form of backchannel signals, which do not constitute instances of turn change. In Example (1), speaker FJCD continues her turn after speaker FLNB has uttered a *ya* ('I see'). This

backchannel is delivered in overlap with part of speaker's FJCD turn, which continues after the backchannel. Therefore, the two parts constitute a single turn[1].

(1) Cross-talk

fjcd_07a: déjame ver. yo la semana que entra tengo una reunión de nueve a once el lunes, /mm/ y después <voy a est> <voy a> voy a estar de viaje. el /uh/ martes miércoles y jueves. o sea el nueve diez y once. [entonces]

flnb_08: [ya].

fjcd_07b: /ah/ <tal vez pode> <podríamos> bueno yo tengo <s> libre el viernes, de once a una.

'Let me see. Next week I have a meeting from 9 to 11 on Monday, mm and then <I'm going to b> <I'm going> I'm going to be away. On uh Tuesday, Wednesday and Thursday. That's the 9th, 10th and 11th.'

'I see.'

'Uh <maybe we ca> <we could> well I have <s> free Friday, from 11 to 1.'

2.2 *Turn-taking signals*

The study of conversational organization has always assumed that there are ways in which speakers communicate the desire to yield, take, or maintain the floor. Sacks et al. (1974) assumed that such signals exist, although they did not discuss any particular signal. Later research has concentrated on a number of different signals: discourse markers, pauses and silence, pitch, and intonation. Of importance are also the syntactic and semantic characteristics of the turn so far, i.e., whether the message can be constructed as a complete one from a syntactic or semantic point of view.

Duncan (1972) proposed that, in every interaction, there are signals that speakers and hearers send to each other in order to indicate their state with regard to the turn. Turn-yielding signals include: intonation (rising or falling pitch); drawl; body motion

---

[1] See Section 3 for a description of transcription conventions.

(termination of hand gesture, relaxation of a tensed hand position); "sociocentric sequences" (fixed expressions such as *or something, you know, but uh*); paralanguage (drops in pitch or loudness); syntax (complete grammatical unit)[2].

Sacks et al. (1974) consider that syntactic information is important. A complete turn-constructional unit is one that can be interpreted as a syntactic unit, whether a sentence, clause, phrase or word. Each one of those units has a component of projectability: the interlocutor knows that the unit is possibly complete from a syntactic point of view.

Content and genre play a role in projectability of a completed unit. Selting (1998) points out that the pre-sequences that introduce narratives are licenses to talk for a relatively long period of time (see also Houtkoop and Mazeland, 1985; Sacks, 1992). Similarly, when telling jokes, the speaker may preface the contribution in a way that makes his or her interlocutor(s) suppress the desire to take the floor (*Have you heard the one about…?*).

Intonation is fundamental in the interpretation of talk. Chafe's (1994) intonation units are defined as basic segments of talk interrupted by the human need to breathe. Intonation units are characterized by changes in pitch (fundamental frequency), duration, intensity, and alternation of talk and silence (pauses). A number of studies examine the pitch characteristics that signal the end of a turn. Beattie et al. (1982) analyzed an interview with Margaret Thatcher, and determined that she signalled the end of a turn when she, however, did not intend to yield the turn. At the interrupted points, she has a fast pitch fall similar to that in her turn-final utterances. This led to frequent interruptions, because the interviewer interpreted the pitch change as a turn-yielding signal. Similarly, in a study involving subjects that either read a transcript or listened to an audio recording of both turn-medial and turn-final utterances, Stephens and Beattie (1986) showed that subjects could identify the turn-final utterances only if presented with an audio recording. Cutler and Pearson (1986) found that there are a few contours that

---

[2] Signals and correlations between signals and turn-taking are further elaborated in Duncan (1973) and Duncan and Fiske (1977).

indicate turn-yielding. Whatever the characteristic, it is clear that phonological information is at play.

Silence and hesitation markers (unfilled and filled pauses) are also signals for interlocutors (Maclay and Oswood, 1959). Beattie (1977) found that people were interrupted more often during a silence (unfilled pause), and that filled pauses tended to follow unfilled pauses. In other words, when a speaker does not produce any talk (unfilled pause), but still wants to hold the floor, a filled pause is produced, to signal the desire to continue talking. Drawl and perceived duration may also play a role in distinguishing turn-medial from turn-final utterances (Stephens and Beattie, 1986).

Ford and Thompson (1996) found that pauses helped identify completed intonation units, the minimal noticeable pause being 0.3 seconds long. However, pauses are not always indicators of an intention to yield the floor. Local and Kelly (1986) proposed that pauses are of two different types: one that signals the intention to keep the floor (a 'holding' silence), and another that indicates that the interlocutor may claim it (a 'trail-off silence'). They particularly examined pauses preceded by a filled pause (*uh, um, so*), following Jefferson (1983). Local and Kelly distinguished the two types of silences phonetically: in holding silences there is a glottal closure after the filled pause, maintained through the silence, and released at the beginning of the following word by the same speaker. Trail-off silences have an out-breathing at the end of the filled pause, which also has a more centralized vowel.

Gaze and gesture are also an important signal in the management of turns. It is clear that gaze can be a turn-allocating mechanism, and that it is an integral part of face-to-face conversation (Beattie, 1979; Lerner, 2003). Kendon (1994; 2002) has shown that gesture fulfills a variety of purposes, among them those typically performed by discourse markers (Kendon, 1995). Speaker and listener movement serve as signals for turn-taking: termination of a hand or arm gesture signals the desire to yield the turn, and continued gesticulation by the speaker acts as a signal to suppress turn-taking by the hearer, similar to the effect of a filled pause (Beattie, 1981; Duncan, 1972). However, since my study

consisted of conversations with no face-to-face interaction, I will not discuss gaze and gesture further.

A combination of features, rather than one individual item, is probably the best indicator of turn boundaries. Duncan and Fiske (1985) present a number of studies that examine the interaction of variables such as gaze, gesture, filled pauses, and the structure of adjacency pairs. Ford and Thompson (1996) studied the interplay of syntactically complete units, intonation and pragmatic closure, and found that, given a syntactically complete unit, it is the combination of intonation (marked fall or marked high rise in pitch at the end of the intonation unit) and pragmatic completion (the unit is interpretable as a complete conversational action) that most often signals a possible turn shift, the complex transition relevance place. Wennerstrom and Siegel (2003) also see turn-taking as a complex process, possible through the interaction of both phonological and syntactic cues. In their study, they examined the interaction of intonation, pauses, and complete syntactic units, concluding that it is a complex interaction of the three that indicates that a speaker's turn has ended, and the floor is open. In some cases, intonation overrides syntax, and it can signal turn continuation despite a syntactic boundary. The intonation pattern with the highest likelihood of indicating a turn shift was the high rise (H-H% in Pierrehumbert's (1980) model), although low rise (L-H%) was also found to indicate turn shift. The most interesting aspect is that not all the high rise utterances were questions from a syntactic point of view. They found that, when longer pauses were produced (0.5 seconds), the current speaker resumed talk. This could be because the opportunity for the other speaker to take the floor, at 0.3 seconds (Ford and Thompson, 1996), had been missed, and the current speaker decided to continue talking. It is also worth noticing that Wennerstrom and Siegel found speaker differences in preference for keeping the floor.

Most of the studies cited above focus on English turn-taking. There are a few studies on prosodic and other signals for turn-taking in languages other than English. For instance, Auer (1996) examines syntactic and prosodic cues in combination in German conversations, and there are a few studies on Japanese (e.g., Hayashi, 1991; Tanaka,

2001). In Spanish, Placencia (1997) points out that in Ecuadorian telephone conversations, closings are very similar to those described by Schegloff and Sacks in English (1973), but that cultural factors account for differences in face-saving strategies (Brown and Levinson, 1978). Hidalgo (1998) examines prosody and its relationship to turn-taking.

In this study, I am also interested in the function of discourse markers as signals in the turn-taking process. By discourse markers I mean a varied group of conjunctions, interjections, filled pauses, adverbs and adverbial phrases, such as *okay, yeah, right, uh-huh, and, so, I mean*. The trouble with discourse markers is that they do not fit one of the three basic types of signals: syntactic, semantic or intonational. They are certainly not prosodic in nature (although they may have their own prosodic characteristics), but it is difficult to say whether they contribute syntactic or semantic information to determine whether the turn is ending, and whether the interlocutor desires to take the turn. Wennerstrom and Siegel (2003), for instance, classify discourse markers together with other syntactic devices that may help in turn-taking.

Discourse markers are generally assumed to signal relations among propositions or among sentences (Fraser, 1999; Knott and Sanders, 1998; Taboada, in press, and many others); they serve to link global and local discourse structure (Redeker, 1990; Schiffrin, 1987), or to indicate a return to a previous topic after a digression (Grosz and Sidner, 1986); they indicate a dispreferred second part in an adjacency pair (Pomerantz, 1984; Schiffrin, 1987); they can serve as acknowledgment tokens (Jefferson, 1984) or backchannel signals (Yngve, 1970); help listeners integrate information in spontaneous talk (Fox Tree and Schrock, 1999); or monitor the interlocutor's comprehension of the speaker's meaning (Fox Tree and Schrock, 2002). Bangerter et al. (2004) examine what they call *project markers (uh-huh, yeah, right, okay)* as signals of transitions between different parts of a telephone conversation. In their analysis, those words have a function at the global level of organization of the conversation (and the task), rather than at the local level of turn-taking.

Sacks et al. (1974: 719) discuss turns that begin with an *appositional beginning*, such as *well, but, and, so*. The appositional beginnings give the speaker some time to think about uttering a complete sentence, and, if there is overlap, make it possible for the interlocutor to capture the full utterance from its "real" beginning (i.e., uttering one of these means: "I'm about to start talking, listen up"). If the previous speaker is not yet listening and missed one of these appositional beginnings, they did not miss any important content. They call them turn-entry devices, or "pre-starts". They propose that these devices be understood as "devices with important turn-organizational uses" (Sacks et al., 1974: 720). These turn-initial (or TCU-initial) signals are also called filled pauses (Beattie, 1977; Maclay and Oswood, 1959).

A few individual discourse markers have been studied as to their role in turn-taking: Schiffrin (1987) reviews the role of turn-initial *well, and, so,* and *but*, showing differences in the content of the turn they start (cooperative, continuative, contrastive). Condon (1986; 2001) has extensively studied the discourse functions of *ok*, among which is the marking of boundaries in decision-making processes (i.e., if not directly a turn-taking device, *ok* helps mark that one portion of the discourse is over, and thus the floor may be open). She also discusses other studies of *ok* that have pointed out its role as closure in phone conversations (Schegloff and Sacks, 1973); as a marker of a topic transition (Schegloff, 1968); or as a marker that the conversation is proceeding as expected, as opposed to *well*, which may signal a dispreferred second part in an adjacency pair (Schiffrin, 1987). A similar function is achieved in Spanish with *pues* (Serrano, 1995) or *bueno* (Briz, 1993).

Redeker (1991) reviews some of Schiffrin's (1987) functions, and adds a few more, among them the function of certain discourse markers in turn-taking: *oh* can elicit a clarification question (thus allocating the turn back to the previous speaker)[3]; *now* and *then* can be used to keep the floor; *I mean* is sometimes a device for starting a turn.

---

[3] On other uses of *oh,* see Local (1996), Heritage (1984), and Schiffrin (1987).

The marker *so* is also thought to be involved in turn-taking, since it serves as a marker of a summary (upshot) of what has previously happened, and can therefore signal the last unit in a turn (Raymond, 2004). Raymond points out that *so* can initiate a unit that is meant to be the upshot, or it can be the upshot itself (without a full unit following *so*).

Ferrara (1997) examines various uses of *anyway*. Among them is the expression by the speaker that he or she wants to regain the floor, after an interruption (although mainly it serves to manage digressions by the speaker).

In summary, the most extensively studied signals for turn taking are intonation, silence, filled pauses and discourse markers. In Section 4 I explore how the last three contribute to maintaining and claiming the floor in a particular corpus of Spanish conversations. Intonation, although an important factor, is not included the present study. The next section provides further information on the corpus used.

3. The corpus

The corpus consists of 60 conversations between dyads of two speakers each, 30 of which are push-to-talk and 30 cross-talk. These terms refer to the recording style. In push-to-talk, speakers have to push the "Enter" key on a computer keyboard in order to gain the floor. In cross-talk, the turn-taking is not impeded by any mechanical means, i.e., it is as natural as in a telephone conversation (Hopper, 1992). Within each group, the conversations are broken down in ten female-female, ten male-male, and ten female-male dialogues. They were chosen to obey, besides the gender balance, two other constraints: that they be approximately the same in length, and that as many different speakers as possible be represented (in the recording of the large corpus, speakers often recorded more than one conversation).

The dialogues were collected by the Interactive Systems Laboratory of Carnegie Mellon University as part of JANUS, a speech-to-speech machine translation project[4]. The two recording styles were tested as part of experiments to determine which type would be easier to process and translate. The speakers were recruited and brought to a lab for the recording. Some of them knew each other beforehand; some of them were introduced at the beginning of the recording session. The instructions explained that the participants have two conflicting agendas (provided to them by the researchers) covering a period of two to four weeks. The participants need to agree on a two hour appointment within that time frame. Further details on the corpus are provided in Taboada (2004).

Speakers were mostly undergraduate and graduate students at Carnegie Mellon University or the University of Pittsburgh. They came from all corners of the Spanish-speaking world: Argentina, Chile, Colombia, Costa Rica, Dominican Republic, Mexico, Peru, Spain, Venezuela, including some raised in the United States (Florida, New York). It is quite possible that there are dialectal variations in the phenomena studied here. My claims are to be interpreted as general with respect to Spanish; further research could detail what aspects are specific to each dialect.

The dialogues were transcribed by members of the JANUS project. The transcripts include a number of conventions introduced by the transcriber, to reflect every sound produced during the conversation. For ease of reading, I have deleted most of them, only leaving some that seemed important: */hm/* and */um/* indicate a hesitation on the part of the speaker. Stretches of talk accompanied by laughter are surrounded by */begin_laugh/* and */end_laugh/*. False starts or repetitions are indicated with < > angled brackets surrounding the material that was repeated or repaired. Overlaps are indicated with square brackets [ ] around the words that are produced at the same time. Backchannels and filled pauses are marked with forward slashes: */uh/*. Pause length is indicated in

---

[4] Thanks to the Interactive Systems Laboratory and to Alex Waibel, its director, for permission to use the corpus. The selection used here is part of a larger corpus of approximately 500 Spanish conversations, and a number of conversations in other languages.

parentheses. Turns are indicated through the initials of the speaker, which always start with either an "f" (for females) or an "m" (males), plus the person's initials, followed by the turn number. Each example indicates whether the source is part of the cross-talk or the push-to-talk corpus.

In addition, transcriber comments include intonation, marked with a comma (,), a period (.) or a question mark (?) at the end of the corresponding section of speech. These markings do not reflect, nor are influenced by, sentence structure. The speaker may have the intonation of a statement when he or she is, in fact, asking a question. He or she may have the falling intonation typical of the end of a sentence (reflected in a period) after a collection of words that do not, in any way, resemble a grammatically correct or complete sentence. For this reason, the word after one of these intonation markers is not usually capitalized, to avoid confusing them with regular periods or question marks.

The translation of Spanish examples is a free translation, rendered one clause at a time. I have tried to capture the overall meaning, rather than translating word-by-word, but staying close to the original, which may make the English translations sound awkward in some cases. When false starts are produced, they sometimes indicate the beginning of a word or phrase that was never completed. It is possible sometimes to guess what the word would have been. In those cases, the English translation reflects a close translation of the interrupted word. In cases where it was not possible to determine the word intended, the false start sound is the same in the translation as in the original.

The conversations are considered instances of a genre, as a purposeful, staged, goal-oriented activity (Martin, 1984), which I have called *scheduling genre* (see also Taboada, 2004). They are instances of talk produced for a very specific purpose, that of setting up an appointment. As a result of their practical purpose, the conversations are staged in particular ways: one speaker proposes a meeting, perhaps also a time; the other speaker replies either with a different time or with their availability for the time proposed. The conversation continues until a day and a time have been set. As a result of their social function, the conversations usually have Opening and Closing stages, and polite devices that will avoid face-threatening acts to the other speaker. The task has a clear structure,

consisting of steps that have to be taken in a certain sequence in order to successfully complete the task.

Tables 1 through 3 provide further information on the corpus. Table 1 lists the number of turns, units and words in both subsets of the corpus. It shows that the cross-talk conversations contain roughly twice as many units and words as the push-to-talk conversations. With regard to turns, the cross-talk conversations have more than three times the number of turns, as compared to the push-to-talk conversations. In other words, speakers in cross-talk conversations say about twice as much as the other speakers, but use almost four times the number of turns.

	Push-to-Talk	Cross-Talk
Turns	248	931
Units	1294	2319
Words	9112	18470

Table 1. Total number of turns, units and words

These tendencies are more clearly presented in Table 2, which provides the average length of turns, units and words, and also the ratio of words to turns (on average).

	Push-to-Talk	Cross-Talk
Average length - turns	8.27	31.03
Average length - units	43.10	77.30
Average length - words	225.93	315.53
Ratio words : turns	27.32	10.17

Table 2. Average lengths per dialogue

Table 3 breaks down the average length according to the speakers' gender. In general, the male-male conversations seem to be longer within the push-to-talk environment. In

cross-talk, it is male-female conversations that are longer under all accounts (turns, units and words). It is difficult to establish whether this is a general tendency, or simply an accident for the ten conversations included in each group.

	Push-to-Talk			Cross-Talk		
	F-F	M-M	M-F	F-F	M-M	M-F
Average length - turns	8.2	8.4	8.2	31.4	26.9	34.8
Average length - units	39.1	49.5	40.7	75.4	69.7	86.8
Average length - words	291.7	342.5	273.1	305.4	302.3	338.9

Table 3. Average lengths per dialogue, by gender. F=female, M=male

4. Turn taking in task-oriented dialogue

The management of floor in specific activities may be slightly different than in spontaneous conversation, the type of interaction that has been described in most studies mentioned so far. For instance, Jones and Thornborrow (2004) describe floor in classroom activities as a result of the activity at hand. They characterize floor "as something people participate in […], rather than "hold"." (Jones and Thornborrow, 2004: 420). In this sense, the concept of floor is similar to Edelsky's (1981) collaboratively developed floor. It is important to note that Edelsky also observed a specific activity, administrative meetings at a university.

In general, the conversations make turn-taking necessary at certain times, because they involve proposals that an interlocutor is, implicitly or not, invited to answer. The presence of a first part in an adjacency pair (Sacks et al., 1974) makes the second part relevant. Lerner discusses this form of *tacit address* in certain action sequences: "Action sequences can play an important part in the selection of a next speaker even when they are not accompanied by an explicit form of addressing. The organization of actions—as sequences of actions—shapes participation; each course of action shapes the opportunities to participate within it. For example, asking a question ordinarily makes an answer especially relevant." (Lerner, 2003: 190). Similarly, in the conversations studied,

16

questions and proposals always invite an answer or a confirmation. This can be done through direct address, but is often achieved implicitly. Obviously, in the push-to-talk conversations, the explicit marker is the pressing of the "Enter" key by the current speaker, signalling the end of his or her turn.

Since strategies for turn-taking are different in each type of corpus, I have divided the strategies in three types, basically the three types of actions that can happen in any conversation. A speaker may (i) yield the turn; or (ii) hold the turn; and an interlocutor may (iii) take the turn. I examine each in the following sections, and for each corpus type.

4.1. *Characteristics of turn yielding*

Turn yielding is the most interesting aspect of the three, since in the push-to-talk conversations there is no possibility of taking the turn (the interlocutor may not take the turn until the current speaker has pressed "Enter"). Turn-yielding, therefore, allows a full comparison across the two types of corpora.

Turn yielding in the push-to-talk conversations is achieved, obviously, by mechanically passing the channel to the interlocutor. But here we are also interested in what happens in the cross-talk conversations when the floor is passed by the current speaker. This is achieved in a variety of ways: pauses, address terms, questions, and tag questions.

Pauses are frequently used as the most obvious signal that the current speaker desires to yield the turn. In (2), speaker FANS proposes a meeting, which projects a second part (an acceptance or a rejection), but makes the desire to yield the turn clearer with a pause.

(2) Cross-talk

    fans_05: oye. <quiero> quiero /ah/ (0.8) tener una conferencia contigo.

    (0.55)

    flxb_06: sí. cómo no. cuándo?

    'Listen. I want I want uh to have a meeting with you.'

'Yes. Of course. When?'

In push-to-talk conversations, pauses are also present. In Example (3), speaker FSNM yields the turn (in the second turn in the example) by pushing a button, but a short pause precedes that pushing. It is, of course, difficult to decide whether the pause is part of the usual turn-taking mechanism, or whether the speaker is simply taking a few seconds to press the turn-yielding key (e.g., trying to find it; or pausing so that the end of her turn is not lost in the transition).

(3) Push-to-talk

ffcs_07: /mm/ el lunes, la verdad que solo tengo una hora en la mañana. pero el martes en la tarde, qué tal?

fsnm_08: pues tengo casi toda la tarde libre a partir de cómo, las doce y media. (1.56)

ffcs_09: (1.13) perfecto. ...

'Mm Monday, the truth is that I only have one hour in the morning. But Tuesday in the afternoon, how's that?'

'I have almost all afternoon free after about twelve thirty.'

'Perfect.'

Pauses were only transcribed if they were at least 0.2 seconds long. Ford and Thompson (1996) report that a length of at least 0.3 seconds is relevant in turn-taking. Therefore, pauses studied here are only those that could be involved in turn-taking. Pause duration was measured using the Praat program, with an additional script that marks pauses in speech[5]. In push-to-talk, pauses at the end of a turn indicate that there was a silence before the speaker pressed the Enter key.

---

[5] Praat is freely available: http://www.fon.hum.uva.nl/praat/. The script can also be downloaded from the following web page: http://www.helsinki.fi/~lennes/praat-scripts/. Many thanks to Mietta Lennes for the script.

In Table 4, I summarize the presence of pauses in the conversations. We can see that they are much more numerous in push-to-talk, especially given that there are fewer turns in those conversations. There is almost one pause per turn, as compared to 27 pauses overall in the 931 turns for the cross-talk data. Striking also is the presence of pauses at turn boundaries. A large number happens at the end of turn, 31.45% of all the pauses present in the push-to-talk data. Presumably, the speaker finishes his or her utterance, and allows for the interlocutor to hear it fully (thus producing a pause) before turning the floor over. However, in some cases, it seems that the interlocutor is not quite ready to talk, as there are also pauses at the beginning of turns. On the other hand, pauses at turn boundaries in cross-talk are minimal, and more frequent at the beginning of a turn, where a speaker may take some time to answer a question that his or her interlocutor has asked.

	Push-to-talk	Cross-talk
Total number of pauses	237	27
At the end of a turn	78 (31.45%)	1 (0.11%)
At the beginning of a turn	36 (14.52%)	6 (0.64%)
Total number of turns	248	931

Table 4. Pauses and turn-taking

Pauses in cross-talk happen most often turn-medially, when speakers provide a chance for their interlocutor to take the floor, which is not taken up. Hopper (1992: 107) describes turn-taking in telephone conversations—which are very similar to the cross-talk conversations analyzed here—as a relay race. A pause is like a runner dropping the baton. "If the runners drop the baton while it is being passed to the next runner, that next runner should retrieve it. If the drop occurs away from such a transition place, the current runner must retrieve it." Such a situation is presented in (4). Speaker FLXB replies to a suggestion of the 22nd with a 'yes' (after a pause), and a repetition of the date. She then

produces a pause, presumably to yield the floor. But the interlocutor does not take up that opportunity, and FLXB repeats the date, making it more explicit that she is free on that day. In the next turn, we understand why speaker FANS did not take the turn at that pause: she wanted more specific information about when on that day FLXB is available.

(4) Cross-talk

fans_11: … qué te parece el veintidós?

flxb_12: (0.65) sí. el veintidós. (0.88) el veintidós está bien. estoy libre.

fans_13: todo el día?

'…What do you think about the 22nd?'

'Yes. The 22nd. The 22nd is good. I am free.'

'All day?'

Address terms are used in very few cases. The speakers may, or may not, have known each other before they came to record the conversations (they were recruited and brought to a lab), but in all cases were introduced to each other, and thus know each other's names. In the following example, the speaker uses a combination of a direct question and her interlocutor's name to yield the turn. Address terms are not necessary, since there is only one interlocutor, and that could explain their scarcity[6].

---

[6] Lerner (2003) also found that address terms are rare, even in multi-party conversation.

(5) Push-to-talk

fmcs_07: … así que /mm/ qué te parece si ya quedamos en juntarnos a la una de la tarde para almorzar, /eh/ no sé donde vos digas, y ya después /eh/ nos queda toda la tarde libre para terminar el proyecto. te parece bien, Miriam?

'… So mm what do you think if we arrange to meet then at 1 p.m. to have lunch, uh I don't know, wherever you want, and then we have the whole afternoon free to finish the project. What do you think, Miriam?'

Direct questions are closely related to address terms. Given that the conversations always involve two people, a direct question is addressed to the only other interlocutor. Thus, even if a question contains no address term, the address term is implied. Example (6) shows a direct address, with the syntactic structure of a question. In fact, speaker MINM realizes early on that this is a question, and does not allow speaker MRRC to finish his utterance, but overlaps as soon as the date (Friday the 23rd) has been mentioned. Example (7) shows a push-to-talk instance of the same phenomenon, with a question (which is actually not answered in the following turn).

(6) Cross-talk

mrrc_09: este, (0.5) qué te parecería, (0.33) <el> (0.54) el viernes veintitrés [de julio.]

minm_10: [viernes veintitrés,] sería perfecto.

'Uh, what about Friday <the> the 23rd [of July.]'

 '[Friday the 23rd,] would be perfect.'

(7) Push-to-talk

mjmg_07: ... el día treinta y uno, (0.95) tengo una clase de dos a cuatro. a ver cómo te va a ti?

msnc_08: bueno. me podrías haber avisado antes. no? ...

'… On the 31st, I have a class from two to four. Let's see how's that with you?'

'Okay. You could have told me before. No?...'

Items projecting a second part in an adjacency pair are not always questions. Sometimes they are suggestions that project an acceptance or a rejection. In (8), the first speaker makes a suggestion to meet at the same place in future occasions. Syntactically, the utterance is a statement, but it finishes with rising intonation. The other speaker realizes the need for an answer, and produces one promptly, although he hesitates after the initial 'yes', which leads to an instance of overlap.

(8) Cross-talk

 fmgg_08: okay. perfecto. <perf> muy bien. <y siempre podríamos> /ah/ las próximas reuniones siempre nos podríamos encontrar en el mismo lugar?

 mmxb_09: sí. sí. <de> [exactamente.]

 fmgg_10: [/eh/ okay.]

 'Okay. Perfect. Perf very good. And <we could always> uh the next meetings we could always meet at the same place?'

 'Yes. Yes. Of exactly.

 'Uh okay.'

Questions may also take the form of tag questions, where confirmation, rather than information, is requested. Sacks et al. (1974: 719) characterize tag questions as exit devices for a turn, or post-completers. They indicate that the turn is complete, and the interlocutor may take the turn. In Example (9), speaker MENM asks for confirmation of the date proposed with a *no?,* a common tag question in Spanish.

(9) Cross-talk

 menm_08: bueno. /eh/ mira. vamos <a> a tratar de encontrar aquí. qué te parece el día nueve?

 mgbm_09: el día nueve de febrero?

menm_10: sí. estamos a dos días de hoy, no?

mgbm_11: /uh_huh/. …

'Okay. Uh look, let's let's try to meet here. How about the 9th?'

'The 9th of February?'

'Yes. That is two days from today, right?'

'Uh-huh. …'

In general, first parts in adjacency pairs seem to be the most common turn-yielding device. As mentioned above, they can be questions or statements asking about or suggesting a meeting time, with a preferred second part being the acceptance of that date, and a rejection as a dispreferred second part. Other pairs include greeting-greeting, goodbye-goodbye, confirmation-confirmation. Table 5 displays the number of first parts in each type of conversation, broken down into questions, tag questions and other. In push-to-talk, the majority of turns (81.85%) involve the initiation of an adjacency pair. In cross-talk, the number is lower (58.97%), but it still accounts for a majority of the turns. That is, speakers yield the turn because they have initiated an adjacency pair that requires a second part, to be produced by the addressee.

	Push-to-talk	Cross-talk
Questions	119	311
Tag questions	4	20
Other first parts	80	218
Total first parts	203 (81.85%)	549 (58.97%)
Total number of turns	248	931

Table 5. Adjacency pairs

Naturally, turns do not consist only of one part in an adjacency pair. Most often a second part starts the turn and a first part finishes it. That is, the sequence is: Speaker A's proposal – Speaker B's rejection + Speaker B's new proposal. Elsewhere I have extensively described the structure of the conversations in those terms (Taboada, 2003, 2004). In (10) we can see a longer example of this succession of adjacency pairs. Speaker FSNM produces a first part, a suggestion of a time to meet. FFCS produces a dispreferred second part, rejecting that date (and giving a reason). She then continues her turn with another first part, a new proposal to meet the following week. FSNM does not directly give a second part to the new proposal, but indirectly rejects it by insisting on meeting on the current week (a new first part). FSNM yields the turn (mechanically, since this is a push-to-talk conversation), and FFCS answers the question with a negative, but does not propose a new date, since she already had suggested moving the date. In fact, turns 4 and 5 are a side sequence, and speaker FFCS's proposal to meet the following week receives a second pair in turn 6. It appears that speakers were keeping track of the side sequence, since it is only turns 4 and 5 that contain a single part of the pair.

(10) Push-to-talk

> fsnm_02: bueno, el lunes para mí será bien como, a la hora de almorzar entre las once y media, por allí, hasta la una y media podremos.

> ffcs_03: no. el lunes no me queda bien. tengo un almuerzo de las doce a las dos. qué te parece si <lo hac> lo dejamos para el próximo lunes.

> fsnm_04: no tienes ninguna hora entre martes y viernes de esa semana?

> ffcs_05: no. no tengo. este <tengo> me voy de viaje por tres días y el viernes tengo clase y dos reuniones.

> fsnm_06: bueno, el lunes, por la mañana será muy bueno para mí pero por la tarde tengo una reunión entre las dos y las cuatro.

> 'Okay, Monday would be good for me, around lunchtime between 11:30, or so, until 1:30 we could.'

'No. Monday is not good for me. I have a lunch between 12 and 2. What do you think if we <do it> leave it until next Monday.'

'Don't you have any time between Tuesday and Friday this week?.'

'No. I don't. Uh <I have> I'm going away for three days and Friday I have a class and two meetings.'

'Okay, on Monday, in the morning would be very good for me, but in the afternoon I have a meeting between 2 and 4.'

The comparison between the two types of conversations shows that silences are proportionately more common in push-to-talk conversations, and that in both types turn-taking seems to follow the structure of adjacency pairs. It is interesting that in the push-to-talk data, a large number of the turns (81.85%) are first parts in an adjacency pair, whereas in cross-talk, although still the majority, first parts are present in 58.97% of the turns. It seems that push-to-talk conversations are more focused on the task, and more compact: most turns contain a task-related element (a first part).

4.2. *Characteristics of turn holding*

In conversation in general, and in task-oriented conversation in particular, speakers may pass the turn to an interlocutor when that interlocutor is not yet ready to take it. As we saw in the previous section, a speaker may utter a first part in an adjacency pair, through a question, a command, or a mention of time availability. Then it is clear that the interlocutor has to answer or otherwise respond to the offer. This happens in both types of conversations. In the push-to-talk data, the passing of the turn is very clear, because the speaker pushes a button that opens the channel to the interlocutor. The other speaker may wish to accept that passing of the turn, but may not be ready to provide a full answer yet. Then he or she holds the floor through a number of devices: silent pauses, filled pauses, and discourse markers.

Silent pauses, that is, periods of time when nobody talks, are the least effective method of holding the floor. A pause may indicate a number of things, among them that

the communication has broken down and needs to be repaired. When the turn is passed by a speaker, and the interlocutor produces a pause, the speaker that passed the turn may feel that something went wrong, and reclaim the turn, maybe providing clarification. Or he or she may not say anything, waiting for the interlocutor to speak. We saw, in Table 4, that pauses at the beginning of a turn are rare: 14.52% of all the pauses in push-to-talk data happened at the beginning of a turn. Recall that in those conversations, the current speaker may hold the floor indefinitely. However, they did not pause very often. One example is given in (11), where speaker MJMG proposes a date, and passes the turn. Speaker MSNC pauses for 1.42 seconds before he narrows down his availability for that date.

(11) Push-to-talk

mjmg_01: (0.96) bueno. vamos a ver. (0.46) /eh/ qué tal el veinticuatro? (0.92) yo puedo cualquier momento, menos de una, a cuatro. (0.67)

msnc_02: (1.42) el veinticuatro yo podría, pero tengo una reunión, desde las diez hasta las doce. tú podrías de diez a doce? (1.12)

'Okay. Let's see. How about the 24th? I can any time, except for from 1, to 4.'

'On the 24th I could, but I have a meeting, from 10 until 12. Could you from 10 to 12?'

Pauses at the beginning of the turn are much less frequent in cross-talk conversation: 6 out of a total of 27 pauses (0.64%) happened at the beginning of a turn, as we saw in the previous section. For example, in (12), speaker FANS pauses at the beginning of her turn (turn number 15), presumably because she is checking her schedule for the day proposed.

(12) Cross-talk:

flxb_12: (0.65) sí. el veintidós. (0.88) el veintidós está bien. estoy libre.

fans_13: todo el día?

flxb_14: todo el día.

fans_15: (0.54) oh yo también. qué bien.

'Yes. The 22nd. The 22nd is good. I'm free.'

'All day?'

'All day.'

'Oh, me too. That's great.'

A filled pause indicates more clearly that the interlocutor, to whom the floor has just been passed, wishes to talk, but is not quite ready to do so. Filled pauses take a number of forms: *eh, ah, mm, uh*. They rarely appear alone, rather being accompanied by a pause, a discourse marker, or both. Example (13) shows a filled pause on its own, *eh*, not only at the beginning of his turn, but also throughout the turn.

(13) Push-to-talk

mrnn_05: /eh/ yo paso por tu oficina /eh/ a las dos de la tarde, <b> porque tengo que /eh/ volver <a c> a casa después de la clase que tengo <en> en la mañana, …

'Uh I'll come by your office uh at two p.m., <b> because I have to uh come back <h> home after the class that I have <in> in the morning, …'

Discourse markers are the most frequent turn-holding device. The definition of discourse marker that I have taken is quite broad. It includes the following: *a ver, vamos a ver* ('let's see'), *oye, oyes* ('listen'), *mira, mirá* ('look'), *ya* ('ok'), *déjame ver* ('let me see'), *pues* ('well/then'), *ay* ('uh oh'), *este* (literally, 'this'), and the English *ok*. There are also combinations of them, such as *mira oye, a ver pues,* or *ya ok*. The distinction between a discourse marker and a filled pause is not always clear. The method that I followed was to classify non-words as filled pauses, and any other adverbs, verbs and conjunctions, as discourse markers.

Byron and Heeman (1998) suggest that discourse markers are more prevalent in task-oriented spoken dialogue: in the TRAINS corpus of task-oriented dialogues that they analyzed, 44.1% of the turns were introduced with a discourse marker (that figure excludes acknowledgements and filled pauses). They found a number of functions were

realized by discourse markers, among them acknowledgements, repairs, and signals about the type of conversational move about to be produced. For example, utterances that summarize previous contributions start with *so*, and utterances that express dissent (a dispreferred second pair) start with *well*. They also found that (preferred) second parts of adjacency pairs did not usually start with a discourse marker. They conclude that discourse markers are used most frequently when there are no strong expectations about the utterance that the speaker is about to make. This seems to be the case in the corpora that I am describing here: discourse markers are used, sometimes in combination with filled pauses, when the speaker him or herself is not sure about what move they need to make next: they need to consider whether they will accept or reject a proposal; or whether they can make a new proposal for a meeting date. An example is presented in (14), where the speaker uses a discourse marker (*la verdad*, 'the truth'), plus a filled pause (*eh*) to reject the date just proposed by the other speaker.

(14) Push-to-talk

fmbo_01: … yo estoy un poco ocupada esta semana porque me voy de viaje. pero (0.43) qué le parece la semana que viene /eh/, <el trece?> el martes trece, a la mañana .

menc_02: la verdad la verdad, /eh/ el trece no es el mejor día para reunirnos para mí.

'… I'm a bit busy this week because I'm going away. But what do you think about next week uh, the 13th? Tuesday the 13th, in the morning.'

'The truth the truth, uh the 13th is not the best day to meet for me.'

In other cases, the discourse marker introduces a request for a clarification, in itself also a dispreferred second part, as in Example (15), from the cross-talk corpus, where both *oh* and *a ver* ('let's see') give the speaker some time to think about the next move: FVNM is not ready to say "yes" to going to a movie until she knows what day is proposed, and as a consequence she requests further information.

(15) Cross-talk

fknh_03: te hablo para decirte de una cita. si quieres ir conmigo al cine.

fvnm_04: oh, a ver. qué día es?

'I'm talking to you to tell you about a meeting. If you want to go with me to the movies.'

'Oh, let's see. What day?'

The discourse marker *bueno* deserves special attention. It is the most common marker, by far, in both types of conversation. Out of the 80 discourse markers present at the beginning of the turn in push-to-talk, 48 are instances of *bueno* (60%). In cross-talk, *bueno* amounts to 53 instances out of a total of 166 discourse markers (32%). *Bueno* is used in isolation, or together with another marker, with a filled pause, or with a pause (*ah bueno, oye bueno*). It is used to signal a dispreferred second part in an adjacency pair (Cortés Rodríguez, 1998), a use similar to that of *well* in English (Schiffrin, 1987). Example (16) illustrates such use of *bueno*: speaker FLNW issues an implicit blanket rejection of all the dates previously proposed by speaker MPNE, and suggests to look at a different week. The rejection is implicit in the *bueno*, and the new proposal seems to be foreshadowed by *mira* ('look').

(16) Push-to-talk

mpne_01: /ehm/ <la> la idea es que como te digo el lunes, en la mañana, martes /eh/ /um/ <de> en la tarde, el miércoles a cualquier hora, y el jueves en la mañana.

flnw_02: bueno. mira. yo creo que vamos a tener que revisar para la semana siguiente porque por lo que conversamos, nuestras agendas están totalmente encontradas. …

'Um <the> the idea is that, as I said, Monday, in the morning, Tuesday uh um <of> in the afternoon, Wednesday anytime, and Thursday in the morning.'

'Well. Look. I think we are going to have to revise for next week because from what we've talked about, our schedules are completely in conflict….'

Example (17) is another instance of *bueno* as a turn-holder. Speaker FMSG has rejected a number of proposals, including the most recent (Friday), and passes the turn. Speaker FDXH does not seem ready to put forth a new proposal, and hesitates with a

number of filled pauses, including *no*, something that sounds close to the French *oh là là*, and then a *bueno*, followed by another filled pause (*ah*), and two new markers combined (*pues mira*), until she finally finds her next available slot (the Wednesday after).

(17) Cross-talk

fmsg_07: …no hombre. se va a poner imposible. el viernes /begin_laugh/ no voy a /end_laugh/ poder. no tengo más que cuatro a cinco. no nos alcanza el tiempo.

fdxh_08: no. <n> ah la lá bueno. ah pues mira. yo no tengo hasta el otro miércoles…

'… No man. It's looking impossible. On Friday I won't be able to. I only have from 4 to 5. That's not enough time.'

'No <n> ah la la well. Ah, well look. I won't have (anything) until the Wednesday after…'

The marker *bueno* has a few other uses not directly related to turn-taking. It can indicate the acceptance of a proposal, and therefore the closing of the conversation, as shown in (18), where it is more appropriately translated as 'okay', instead of 'well', as we have seen in the previous examples.

(18) Push-to-talk

ffcs_07: /mm/ el lunes, la verdad que solo tengo una hora en la mañana. pero el martes en la tarde, qué tal?

fsnm_08: pues tengo casi toda la tarde libre a partir de como, las doce y media. (1.56)

ffcs_09: (1.13) perfecto. porque no nos reunimos <a la una> /ah/ de una a las tres. qué tal? (1.7)

fsnm_10: (0.92) será muy bueno esa hora, entonces te veré allí.

ffcs_11: bueno. está bien. te veo el martes a la una. hasta luego.

'Mm, on Monday the fact is that I only have one hour in the morning. But Tuesday in the afternoon, how's that?'

'Well, I have almost all afternoon free, from about 12:30 on.'

'Perfect. Why don't we meet <at one> uh from one to three. How's that?'

30

'It'll be very good at that time, then I'll see you there.'

'Okay. That's good. I'll see you on Tuesday at one. See you later.'

Repetitions also serve as turn-holders. In (19), speaker FANS repeats, in the last turn of the example, the date just proposed (the 17th), with rising intonation. This could be both a request for a confirmation and a turn holder. The request for confirmation hypothesis is supported by the fact that there is a pause after *el diecisiete?*, but the turn holder hypothesis is possible, since FANS continues to talk without having received confirmation. She may just have said something to hold the turn, and to indicate that she is considering that date.

(19) Cross-talk:

    fans_09: y a ver. qué otro día yo puedo.

    flxb_10: (0.64) el diecisiete puedes?

    fans_11: el diecisiete? (1.59) no puedo el diecisiete. (1.8) <qu> qué te parece el veintidós?

    'And let's see what other day I can.'

    'Can you on the 17th?'

    'The 17th? I can't on the 17th. <Wh> What do you think of the 22nd?'

Pauses, filled pauses, and discourse markers are also used in combination. In Example (20), speaker MFMM starts with a marker (*a ver*), and then continues with a filled pause and an address term, before pointing out a problem with the time just proposed by his interlocutor.

(20) Push-to-talk

    mfmm_08: a ver, /eh/ Octavio. creo que las dos horas tienen que ser seguidas. va a ser una reunión bastante larga, tiene que ser de dos horas y las dos horas seguidas. …

    'Let's see, uh Octavio. I think that the two hours have to be contiguous. It's going to be a pretty long meeting, it has to be two hours and the two hours contiguous. …'

Table 6 summarizes the numbers of filled pauses and discourse markers in each of the corpora, and their averages per conversation. This represents only instances that occur at the beginning of the turn.

	Push-to-talk	Cross-talk
Filled pauses	24	108
Discourse markers	80	166
Filled pauses per conversation	0.8	3.6
Discourse markers per conversation	2.6	5.5

Table 6. Filled pauses and discourse markers

Although it is difficult to compare the numbers (given that cross-talk conversations contain more turns and more words), we can conclude that, on average, each push-to-talk conversation contained fewer filled pauses and fewer discourse markers. In other words, the speakers completed the same task using fewer filled pauses and discourse markers in the push-to-talk setting. One conclusion of this data is that those features are necessary when managing spontaneous conversation, but they drop in frequency once the pressure of holding the turn disappears.

4.3. *Characteristics of turn taking*

Turn-taking in the push-to-talk conversations does not happen freely, but is controlled by the interlocutor: a speaker can take the turn only if the interlocutor yields it. Once the turn has been passed, we are in a turn-holding situation. It is only in the cross-talk conversations that we can observe actual instances of turn taking, that is, a speaker taking the floor from the current speaker. In Sacks et al.'s (1974) terms, the speaker is self-selecting at the transition-relevance place. Since there is no possible comparison in this

respect between the two corpora, I will not have much to say about how turn-taking happens, apart from some observations about overlapping.

Overlapping is the most obvious instance of self-selection by an interlocutor, since it indicates that the current speaker is not ready to yield the turn. In the cross-talk data, 22% of the turns contain some overlapping or simultaneous talk (204 out of 931 turns). In most cases, the simultaneous talk is clearly not an instance of trying to take the floor, but merely a backchannel, as in (21), where speaker MLGZ produces a backchannel that signals understanding that Monday is not good (*ah*), but then lets speaker MLPM finish his turn before he asks the following question ('and when are you back').

(21) Cross-talk

mlpm_03: este el lunes estoy de viaje. [el lunes,] <me> muy complicado para mí.

mlgz_04: [ah.] y cuándo volvés?

'Eh Monday I'm away. [Monday,] <me> very busy for me.'

'[ah] and when are you back?'

In other cases, it is more plausible to think that the interlocutor was trying to take the floor, sometimes to preclude discussion of a date, or to make the interlocutor stop when a date has been proposed, so that the self-selecting speaker gets the chance to consult his or her calendar. In (22), speaker FDXH suggests Friday, and ends her question with rising intonation. At that point, speaker FMSG repeats 'Friday', also with rising intonation, ready to take the floor and check her schedule. But speaker FDXH has not finished her turn, and repeats the Friday proposal, specifying that she is free then. The overlapping talk happens at the point where FDXH repeats *el viernes*, because FMSG was prompt to take the floor at that point. In fact, FMSG could naturally have expected her interlocutor's turn to end at the point where FDXH produced a question. According to Schegloff (1988: 141), "if a turn has several components (that is, turn-constructional units) in it, one of which is a question, the question is almost always the last of them, for on its completion, the question will ordinarily have made it someone else's turn to talk.".

(22) Cross-talk

fdxh_04: bueno. qué te parece /eh/ <ma> el viernes? [el viernes] tengo libre.

fmsg_05: [el viernes?] el viernes. ah caray. déjame ver. este, tengo una reunión. con este David ay ya no me acuerdo cómo se apellida. /begin_laugh/ <d> de /end_laugh/ diez a doce.

'Okay. What do you think uh <ma> on Friday? [Friday] I'm free.'

'[Friday?] Friday. Oh jeez. Let me see. Um, I have a meeting. With this David, uh I don't remember his last name. <f> from ten to twelve.'

5. Discussion

I have presented an analysis of turn-taking strategies in two different sets of task-oriented conversations. The comparison between the two is informative because they were recorded in exactly the same circumstances, with the exception of the turn-taking mechanisms: speakers in one group of conversations controlled the turn mechanically, resulting in one-way communication.

The conversations have been examined in terms of three characteristics: turn yielding, turn holding and turn taking. The two types are different in terms of turn yielding and turn holding. In both cases, pauses play a role. Pauses are more frequent at the ends and beginnings of turns in the push-to-talk data, maybe a reflection of the mechanical intervention. On the other hand, cross-talk, that is, spontaneous conversations, have a higher number of filled pauses, probably as a result of the need to show that the turn is not being yielded.

The most significant differences overall are, in push-to-talk, the lower numbers of pauses, filled pauses and discourse markers; in cross-talk, the presence of overlap. As it has been pointed out before, those are common characteristics of spontaneous conversation. In general, pauses, if present, are very brief, and overlap exists, but it is also brief. Filled pauses and discourse markers are used to manage the interaction, and to hold the turn. When we control the turn-taking through mechanical means, the typical features of turn-taking decrease in frequency.

The conclusions that we can draw from this data are potentially limited to the data itself. Neither type of conversation was completely spontaneous. Even the cross-talk data was recorded in a lab, in somewhat artificial conditions. Most conversation analysis practitioners would frown upon such data. I feel that we cannot discard data because it does not conform to some ideal of what "spontaneous" means. At the very least, we can conclude that these are perfect instances of conversations produced in a particular setting: the recording laboratory. Since we have a large number of conversations, and since they were all recorded under similar circumstances, I believe that generalizations about their characteristics are possible.

The research into turn-taking signals is useful in itself: we want to understand how conversational participants signal to each other that they desire to yield or take the turn. But it has many other applications, one of them the design of intelligent conversational agents. Researchers in Artificial Intelligence have been designing systems that can interact with a user, provide directions or information. One example is a conversational agent developed at MIT that provides a virtual tour of real estate for sale. Bickmore and Cassell (2005) found that the virtual real estate agent needs to know when and how to yield and take the turn, using not only verbal cues, but also gesture and posture. The nonverbal cues were also important in building trust with the virtual agent.

References

Auer, Peter (1996) On the prosody and syntax of turn-continuations. In E. Couper-Kuhlen and M. Selting (eds.), *Prosody in Conversation: Interactional studies*. Cambridge: Cambridge University Press, 57-100.

Bangerter, Adrian, Herbert H. Clark and Anna R. Katz (2004) Navigating joint projects in telephone conversations. *Discourse Processes* 37.1: 1-23.

Beattie, Geoffrey (1977) The dynamics of interruption and the filled pause. *The British journal of Social and Clinical Psychology* 16.3: 283-284.

Beattie, Geoffrey (1979) Planning units in spontaneous speech: Some evidence from hesitation in speech and speaker gaze direction in conversation. *Linguistics* 17: 61-78.

Beattie, Geoffrey (1981) The regulation of speaker turns in face-to-face conversation: Some implications for conversation in sound-only communication channels. *Semiotica* 34.1-2: 55-70.

Beattie, Geoffrey, Anne Cutler and Mark Pearson (1982) Why is Mrs. Thatcher interrupted so often? *Nature* 300.23: 744-747.

Bickmore, Timothy W. and Justine Cassell (2005) Social dialogue with embodied conversational agents. In J. van Kuppevelt, L. Dybkjaer and N. Bernsen (eds.), *Advances in Natural Multimodal Dialogue Systems*. Berlin: Springer.

Briz, Antonio (1993) Los conectores pragmáticos en español coloquial (I): Su papel argumentativo. *Contextos* 11.21-22: 145-188.

Brown, Penelope and Stephen Levinson (1978) Universals in language use: Politeness phenomena. In E. N. Goody (ed.), *Questions and Politeness: Strategies in Social Interaction*. Cambridge: Cambridge University Press, 56-289.

Byron, Donna K. and Peter A. Heeman (1998) Identifying discourse markers in spoken dialog, *AAAI Spring Symposium on Applying Machine Learning and Discourse Processing*. Stanford, CA.

Chafe, Wallace (1994) *Discourse, Consciousness and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.

Condon, Sherri L. (1986) The discourse functions of OK. *Semiotica* 60: 73-101.

Condon, Sherri L. (2001) Discourse *ok* revisited: Default organization in verbal interaction. *Journal of Pragmatics* 33: 491-513.

Cortés Rodríguez, Luis (1998) Marcadores del discurso y análisis cuantitativo. In M. A. Martín Zorraquino and E. Montolío Durán (eds.), *Los marcadores del discurso: Teoría y análisis*. Madrid: Arco, 143-160.

Cutler, Anne and Mark Pearson (1986) On the analysis of prosodic turn-taking cues. In C. Johns-Lewis (ed.), *Intonation in Discourse*. San Diego, CA: College Hill, 139-156.

Duncan, Starkey (1972) Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology* 23: 283-292.

Duncan, Starkey (1973) Toward a grammar for dyadic conversation. *Semiotica* 9.1: 29-46.

Duncan, Starkey and Donald W. Fiske (1977) *Face-to-Face Interaction: Research, Methods, and Theory*. Hillsdale, NJ: Lawrence Erlbaum.

Duncan, Starkey and Donald W. Fiske (1985) *Interaction Structure and Strategy*. Cambridge: Cambridge University Press.

Edelsky, Carole (1981) Who's got the floor? *Language in Society* 10: 383-421.

Ferrara, Kathleen (1997) Form and function of the discourse marker 'anyway': Implications for discourse analysis. *Linguistics* 35: 343-378.

Ford, Cecilia, Barbara A. Fox and Sandra A. Thompson (1996) Practices in the construction of turns: The 'TCU' revisited. *Pragmatics* 6.3: 427-454.

Ford, Cecilia and Sandra A Thompson (1996) Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. Schegloff and S. A. Thompson (eds.), *Interaction and Grammar*. Cambridge: Cambridge University Press, 134-184.

Fox Tree, Jean E. and Josef C. Schrock (1999) Discourse markers in spontaneous speech: Oh what a difference an oh makes. *Journal of Memory and Language* 40: 280-295.

Fox Tree, Jean E. and Josef C. Schrock (2002) Basic meanings of *you know* and *I mean*. *Journal of Pragmatics* 34: 727-747.

Fraser, Bruce (1999) What are discourse markers? *Journal of Pragmatics* 31: 931-952.

Furo, Hiroko (2001) *Turn-Taking in English and Japanese: Projectability in Grammar, Intonation, and Semantics*. New York: Routledge.

Goffman, Erving (1981) *Forms of Talk*. Pennsylvania: University of Pennsylvania Press.

Goodwin, Charles (1981) *Conversational Organization: Interaction between Speakers and Hearers*. New York: Academic Press.

Grosz, Barbara J. and Candace L. Sidner (1986) Attention, intentions, and the structure of discourse. *Computational Linguistics* 12.3: 175-204.

Hayashi, R. (1991) Floor structure of English and Japanese conversation. *Journal of Pragmatics* 16: 1-30.

Heritage, John (1984) A change-of-state token and aspects of its sequential placement. In J. M. Atkinson and J. Heritage (eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, 299-345.

Hidalgo, Antonio (1998) Alternancia de turnos y conversación. Sobre el papel regulador de los suprasegmentos en el habla simultánea. *Lingüística Española Actual* 20.2: 217-238.

Hopper, Robert (1992) *Telephone Conversation*. Bloomington, IN: Indiana University Press.

Houtkoop, Hanneke and Harrie Mazeland (1985) Turns and discourse units in everyday conversation. *Journal of Pragmatics* 9: 595-619.

Jaffe, Joseph and Stanley Feldstein (1970) *Rhythms of Dialogue*. New York: Academic Press.

Jefferson, Gail (1983) On a failed hypothesis: 'Conjunctionals' as overlap-vulnerable. *Tilburg Papers in Language and Literature* 28: 29-33.

Jefferson, Gail (1984) Notes on a systematic deployment of the acknowledgment tokens "yeah" and "mm hm". *Papers in Linguistics* 17: 197-216.

Jones, Rod and Joanna Thornborrow (2004) Floors, talk and the organization of classroom activities. *Language in Society* 33: 399-423.

Kendon, Adam (1994) Do gestures communicate?: A review. *Research on Language and Social Interaction* 27.3: 175-200.

Kendon, Adam (1995) Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics* 23.3: 247-279.

Kendon, Adam (2002) Some uses of the head shake. *Gesture* 2.2: 147-182.

Knott, Alistair and Ted Sanders (1998) The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics* 30: 135-175.

Lerner, Gene H. (2003) Selecting next speaker: The context-sensitive operation of a context-free organization. *Language in Society* 32.2: 177-201.

Local, John (1996) Conversational phonetics: Some aspects of news receipts in everyday talk. In E. Couper-Kuhlen and M. Selting (eds.), *Prosody in conversation*. Cambridge: Cambridge University Press, 177-230.

Local, John and John Kelly (1986) Projection and "silences": Notes on phonetic and conversational structure. *Human Studies* 9: 185-204.

Maclay, Howard and Charles E. Oswood (1959) Hesitation phenomena in spontaneous English speech. *Word* 15: 19-44.

Martin, James R. (1984) Language, register and genre. In F. Christie (ed.), *Children Writing: Reader*. Geelong, Victoria: Deakin University Press, 21-30.

Pierrehumbert, Janet (1980) *The Phonology and Phonetics of English intonation.* Ph.D. dissertation, MIT, Cambridge, Mass.

Placencia, María Elena (1997) Opening up closings---the Ecuadorian way. *Text* 17.1: 53-81.

Pomerantz, Anita (1984) Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson and J. Heritage (eds.), *Structures of Social Interaction: Studies in Conversation Analysis.* Cambridge: Cambridge University Press, 57-101.

Raymond, Geoffrey (2004) Prompting action: The stand-alone "so" in ordinary conversation. *Research on Language and Social Interaction* 37.2: 185-218.

Redeker, Gisela (1990) Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14: 367-381.

Redeker, Gisela (1991) Review article: Linguistic markers of linguistic structure. *Linguistics* 29.6: 1139-1172.

Sacks, Harvey (1992) *Lectures on Conversation* (Gail Jefferson ed. Vol. II). London: Sage.

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 (1968) Sequencing in conversational openings. *American Anthropologist* 70: 1075-1095.

Schegloff, Emmanuel (1982) Discourse as an interactional achievement: some uses of uh huh and other things that come between sentences. In D. Tannen (ed.), *Analyzing Discourse: Text and Talk. Georgetown University Roundtable on Languages and Linguistics.* Washington, DC: Georgetown University Press, 71-93.

Schegloff, Emmanuel (1988) Discourse as an interactional achievement II: An exercise in Conversation Analysis. In D. Tannen (ed.), *Linguistics in Context: Connecting Observations and Understanding.* Norwood, NJ: Ablex, 135-158.

Schegloff, Emmanuel and Harvey Sacks (1973) Opening up closings. *Semiotica* 8: 289-327.

Schiffrin, Deborah (1987) *Discourse Markers.* Cambridge: Cambridge University Press.

Selting, Margret (1998) TCUs and TRPs: The construction of units in conversational talk. *InLiSt - Interaction and Linguistic Structures* 4: http://inlist.uni-konstanz.de/issues/4/index.htm.

Selting, Margret (2000) The construction of units in conversational talk. *Language in Society* 29: 477-517.

Serrano, María José (1995) El uso de *la verdad* y *pues* como marcadores discursivos de respuesta. *Español Actual* 64: 5-16.

Stephens, Jane and Geoffrey Beattie (1986) On judging the ends of speaker turns in conversation. *Journal of Language and Social Psychology* 5.2: 119-134.

Taboada, Maite (2003) Modeling task-oriented dialogue. *Computers and the Humanities* 37.4: 431-454.

Taboada, Maite (2004) *Building Coherence and Cohesion: Task-Oriented Dialogue in English and Spanish.* Amsterdam and Philadelphia: John Benjamins.

Taboada, Maite (in press) Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics.*

Tanaka, Hiroko (2001) Adverbials for turn projection in Japanese: Toward a demystification of the "telepathic" mode of communication. *Language in Society* 30.4: 559-587.

Wennerstrom, Ann and Andrew F. Siegel (2003) Keeping the floor in multiparty conversations: Intonation, syntax, and pause. *Discourse Processes* 36.2: 77-107.

Yngve, Victor H. (1970) On getting a word in edgewise. In *Papers from the Sixth Regional Meeting of the Chicago Linguistics Society*. Chicago: University of Chicago, 567-577.