

## **A STUDY ON SPEECH RATE AS A PROSODIC FEATURE IN SPONTANEOUS NARRATIVE**

Miguel OLIVEIRA JÚNIOR.\*

- **ABSTRACT:** Speech rate is examined in this paper as a prosodic feature employed in the signaling of spontaneous narrative structure. Assuming that narratives have a structural system in itself, and that interactants mark their moves and their more global activities in order to make them unambiguous (JEFFERSON, 1978; SACKS, 1972), the present paper examines speech rate phenomena, from an acoustic-experimental approach, in 17 spontaneous narratives, using one of the most influential models for narrative analysis - the Labovian Evaluative Model (LABOV, 1972) - as framework for the analysis. The prosodic variable under investigation is analyzed on two different levels: at specific points in the narratives corresponding to section boundaries (local level), and within different sections in the narratives as a whole (global level). The results indicate that speech rate operates exclusively on the global level, by generating a cyclical pattern of varying rates corresponding to the individual, linear sections that make up narrative texts. Speech rate does not characterize narrative sections and is not manipulated on the local level in order to mark narrative boundaries.
- **KEYWORDS:** Spontaneous narrative. Prosody. Discourse structure. Speech rate.

### **Introduction**

The degree of syntactic, semantic, and/or pragmatic cohesiveness between words in an utterance determines whether they belong together to a larger linguistic constituent or not. To the same extent, utterances bear different sorts of relations with other utterances in an even larger linguistic constituent that, when grouped together, form what is generally referred to as a “discourse”.<sup>1</sup> In this view, discourse is considered to be a structure composed by hierarchically arranged entities that preserve a similar orientation. In written language, these entities are sometimes called “paragraphs.” They are often signaled by typographic means, such as an indent line at the beginning and an incomplete line at the end (which

---

\* UFAL – Universidade Federal de Alagoas. Faculdade de Letras – Maceió – AL – Brasil. 57309-005 – miguel@fale.ufal.br

<sup>1</sup> Leech and Short (1981, p.209) “[...]distinguish “discourse” from “text” in terms of the functions each of these concepts convey: the later is regarded as a “message in its auditory or visual medium,” while the former is viewed as an “interpersonal activity.” These definitions resemble the common – and misleading – discrimination of linguistic communication between “monologue” and “dialogue.” In this paper, the words “discourse” and “text” will be used interchangeably.”

may be absent in cases where the end of the paragraph coincide with the end of the line). Spoken discourse also presents such macro-structures, which are referred to as “discourse segments” (PASSONNEAU; LITMAN, 1993), “topics”, “information units” (SWERTS; GELUYKENS, 1994), and even “paragraphs” (LEHISTE, 1975). These units are marked in speech by the use of different linguistic phenomena, such as anaphora (GROSZ; SIDNER, 1986), cue phrases (PASSONNEAU; LITMAN, 1993), discourse markers (SCHIFFRIN, 1987), reference (GROSZ; SIDNER, 1986) and tense (HWANG; SCHUBERT, 1992).

One of the most important structuring or demarcative devices in spoken discourse is prosody. Variation in pitch range (BROWN; CURRIE; KENWORTHY, 1980; HIRSCHBERG; GROSZ, 1992; SILVERMAN, 1987; SWERTS 1997; among others), pausal duration (SWERTS; GELUYKENS, 1994; GROSZ; HIRCHBERG, 1993; COLLIER, PIYPER; SANDERMAN, 1993; among others), speech rate (LEHISTE, 1982; KOOPMANS-VAN BEINUM; VAN DONZEL, 1996; FON, 1999; SELTING, 1992), and amplitude (BROWN; CURRIE; KENWORTHY, 1980; HIRSCHBERG; GROSZ, 1992; GROSZ; HIRCHBERG, 1993) have all been studied, with some success, as potential correlates of discourse structure in speech.

Independent of any prosodic evidence, some discourse types (or genres) are considered to have an internal structure that can be observed solely by taking into account the content of their constituents. Narratives, for example, are thought to be composed of semantically independent segments (sections or units) that can be easily recognized.

For that reason, narratives are thought to have an underlying grammar that can be used to describe and generate narrative discourse. (GLENN, 1978). Several approaches to describing this underlying grammar (or model) of narrative discourse have been proposed. Literary theorists, for instance, have used structuralist or generative models of language to create models of how stories are constructed and what plots are like. (BARTHES, 1975; GENETTE, 1980; PRINCE, 1982). Story grammarians have attempted to predict universal processing regularities in narratives in order to explicate implicit nonlinguistic knowledge elements necessary for story processing. (BLACK; WLILENSKY 1979; MANDLER; JOHNSON 1977; RUMELHART 1980). Conversational analysts have considered the mutual activity of storytelling as a structural system in itself, by assuming that interactants mark their moves and their more global activities in order to make them unambiguous. (JEFFERSON, 1978; RYAVE, 1978; SACKS, 1972).

One of the most influential narrative models in linguistics research is that of Labov and Waletzky (1967) and Labov (1972): the Evaluative Model. In this model, oral narratives are shown to be bounded discourse units that can be segmented according to their informational function. Labov (1972), in expanding

on his previous work with Waletzky, proposes six elements in the structure of a well-formed narrative: (1) abstract; (2) orientation; (3) complicating action; (4) evaluation; (5) resolution; (6) coda. These sections are listed in their usual order of occurrence (except for the “evaluation,” which may be found in various forms throughout the narrative).

The **Abstract** initiates the narrative by summarizing the point of the story a teller intends to follow, or by providing a statement of a general proposal, which the story itself will exemplify. The **Orientation** usually gives detailed information about the time, characters, situation and place where the event(s) occurred — the background, which, the narrator believes, the audience requires to understand the story. The **Complication** consists of a series of narrative clauses that answers the question: “then, what happened?” It is the backbone of the story and builds up to its climax. The **Result** contains the resolution to a conflict in the narrative. It usually contains free clauses, which began the complicating action. The **Coda** signals the “sealing off” of a story, by returning the listeners to the present moment. The **Evaluation** consists of all the possible means employed by a teller to situate and support the point, tellability or reportability of his/her story. It may take a multitude of forms and surface at almost any point in the telling, although it is often clustered around the climactic point of the action, just before the resolution.

The purpose of this paper is to investigate whether a specific acoustical-temporal prosodic feature, speech rate, is employed as a cue for narrative segmentation. Following a tradition in this type of study (ARCHAKIS; PAPAZACHARIOU, 2008; ATTARDO; PICKERING, 2011; FERRÉ, 2009; PICKERING et. al., 2009), the Labovian Evaluative Model (LABOV, 1972) will be used as framework for the analysis.

### **Speech rate as a segmentation cue in narrative texts**

Variation in speech rate is sometimes regarded as a supplementary prosodic cue employed in the segmentation of discourse. Koopmans-van Beinum and Van Donzel (1996), for example, demonstrated that speakers often slow down at the start of a new paragraph and speed up at the end of paragraphs, in personal comments and additions. After conducting measurements of the average syllable duration (ASD) of speech samples derived from spontaneous and read-aloud narratives by eight speakers of Standard Dutch, the authors found a relatively large number of cases in which peak ASD-values co-occurred with discourse markers, such as ‘and then’. Since in most of the cases these markers indicated the beginning of a new topic, they concluded that there exists a relationship between discourse structure and speech variability.

Through the analysis of a narrative told in the course of a conversation in German, Selting (1992) verified that the distribution of accents within the complication action of the narrative was roughly placed in equal distances, resulting in intonation units shorter than usual. The consequence of this was a sort of “speeding up” characterizing the foreground information of the story. This would, once again, suggest that changes in speech rate can be manipulated in order to contextualize what is said. (JHMANN, 1992).

This “manipulation” of speech rate is also exemplified by Fon (1999). In her study, speech rate is shown to fluctuate within narratives and most speakers make use of a strategy so that a one-to-one relationship between the structure of different narrative types and rate cycle could be observed. According to Fon (1999), this would reflect a predisposition of speakers to plan their own speech in order to accommodate complete discourse units. She also suggests that if this high correlation between speech rate and story parts is regularly employed by speakers as a cue for narrative segmentation, it would be very likely that listeners use this cue as a way of processing the incoming signal.

Grosz and Hirschberg (1992) and Hirschberg and Grosz (1992) demonstrated that, at the local level of discourse, parenthetical phrases are characterized by a higher speech rate (6.05 syllables per second, as opposed to 5.04 syllables per second in their data). They also found that rate, along with other acoustic-prosodic features, is responsible for the categorization of attributive tags and phrases beginning direct quotations. However, according to their analysis, rate was not found to have a major influence on the global level of discourse segmentation.

It should be pointed out that the study of speech rate is closely related to that of pause in speech in many ways. There are reasons to believe that both pausological research and studies on speech rate are branches of just one major area of research: pausology (O’CONNELL; KOWAL, 1983). Some authors argue that the major determiner of speech rate is not speech *per se* but rather pause time. Goldman-Eisler (1956), for example, demonstrated that speech rate is more closely related to variation in length and frequency of unfilled pauses. According to her findings, the variability of speech rate is a function of the high degree of variability in the time which speakers spend hesitating between sequences of actual speech (see also GOLDMAN-EISLER, 1968). Sabin (1976) notes from his data that the variation in speech rate is attributed to length of pauses in 74% of the cases and frequency of pauses in 69% of the cases. The same findings appear in Sabin et al. (1979) and Grosjean (1980), indicating a strong relationship between these two temporal phenomena.

As previous studies demonstrated (OLIVEIRA JÚNIOR, 2002a), pausal phenomena can be manipulated in narratives in order to make the structure of this type of discourse transparent to the audience. Since speech rate is often

considered to be intrinsically related to pausal phenomena, it is expected that storytellers make use of variation in speech rate to structure their narratives the same way they do with pause.

Speech rate in this paper will be analyzed on both the global and the local levels. On one hand, it is hypothesized that narrative sections are characterized by the use of different rates and that this variation forms a temporal cycle similar to the one found for pausing. On the other hand, it is expected, on the basis of the high correlation between speech rate and pausal phenomena reported in the literature (OLIVEIRA JÚNIOR, 2002b), that the difference in speech rate between two adjacent intonation units will be higher when it coincides with a narrative boundary. This difference in rate would be used to indicate that a new section is about to begin, serving thus as a cue for narrative segmentation as well.

## **Material and methods**

A total of eight people (six women and two men) participated as informants in this study. Most were graduate students at the time of the data collection, ranging from 25 to 37 years of age. No one knew the details of the study or the researcher's specific area of study.

The subjects were asked to talk freely on any topic from a list of 28 possible topics and the researcher only acted as an interviewer, stimulating the talk and providing some feedback responses; each participant was presented with a different version of the list, which contained exactly the same topics, but enumerated in a random order, as to avoid the recursion of a specific subject. They were instructed to pick any topic in any order they wished.<sup>2</sup>

Recordings were made in a sound-treated room using a professional cassette recorder (Marantz PMD201) and a unidirectional dynamic microphone (Genexxa Intertan 33-984 DCA), positioned about 15-30 cm from the participants' mouth. The total duration of each interview ranged from 45 minutes to 62 minutes. In general, the participants behaved uneasily in the first few minutes of the recording session, a common reaction in situations like this. (WOLFSON, 1976). However, they all appeared to be relaxed after a period of approximately ten minutes and spoke with a high degree of spontaneity. The fact that most of the participants were acquainted with the researcher contributed greatly to the high degree of spontaneity in most of the recordings. All the narratives that were selected for this study were extracted after a minimum period of ten minutes from the beginning of the recording. They appeared naturally in the discourse, most of the time as an illustration of a given argument or topic. The participants were not asked

---

<sup>2</sup> This procedure yields what Wolfson (1979) calls "spontaneous interviews".

to tell stories, nor was it suggested in any way that narratives were to be told. Nevertheless, most speakers naturally told at least one narrative.

The definition of what characterizes “spontaneous speech” and the elicitation methods that are employed to collect this speech style are surrounded by controversies. Speech is usually labeled as being either spontaneous or read. This polar classification is misleading, because it does not take into account the *continuum* that lies in between. (FUJISAKI, 1997; LAAN, 1997). Because most scholars prefer to deal with a polar (rather than a gradual) classification of speech, determining what in fact is “spontaneous speech” has been a very difficult task.

In general, speech is designated as spontaneous in terms of a series of linguistic aspects that are often present in unprepared utterances (or utterances prepared to a minimum degree), such as the occurrence of disfluencies and repairs, the frequent use of abbreviations, a more relaxed syntax, and the incidence of “fillers,” like ‘ers’ and ‘ahs’. Studies have demonstrated that listeners are able to make a clear judgment about speech using a binary classification (spontaneous versus read). (LAAN 1997; LEVIN; SCHAFFER; SNOW, 1982). This would suggest the validity of such a classification.

This is not a new debate. It goes back to Labov’s (1971) sociolinguistic discussion of the *observer’s paradox*. According to Labov (1971), there are a couple of elicitation techniques that can be used to overcome the *observer’s paradox* in fieldwork. One of the most successful and, for that reason, widely used in (socio) linguistic investigation, is the prompting of narratives with open-ended questions. It is generally assumed that narratives are excellent examples of fluent speech, because when one is telling a story, s/he usually overlooks the conversational setting and transports her/himself to the story world, resulting in a closer attention to the content rather than to the form of the talk. The way narratives are elicited, however, may have a very significant impact on the resulting product.

When a speaker decides to tell a story during a conversation, without being prompted to do so, s/he makes explicit her/his judgment of what s/he believes the audience would find worthwhile enough to justify relinquishing their rights to the conversational floor. This obviously puts much more responsibility on the narrator, as s/he, out of necessity, will have to make an effort to demonstrate that the narrative s/he is telling is not only relevant to the talk, but worthy of being heard. Narratives that are told as an answer to a request may have a completely different characteristic than those occurring naturally in the conversation.

In contrast to unprompted narratives, the most immediate point of elicited stories is to respond to a question. In principle, it would not be expected from the narrator anything more than that. It would not be surprising, then, that elicited narratives were qualitatively different from spontaneous, non-elicited

stories. According to Wolfson (1979), people know which rules of speaking are appropriate for interviews as speech events. That would be the reason why some narratives, when elicited, assume the form of summaries: they are often short, to the point and display very few details, as answers to questions in an interview are supposed to be.

The present study considers spontaneous, non-elicited narratives as a legitimate sample of “real” spontaneous speech.

From all interviews, a total of 25 texts were initially identified as narratives. Of these, only 17 texts were selected to form the main database of the present study. The selection of the narratives was made by taking into account the following criteria:

- (1) The basic definition of narrative, as proposed by Labov (1972). Texts that could be intuitively classified as narratives, but did not show a sequence of “discrete, temporally and non-randomly ordered units,” were discarded from the study. Three texts that did not fall into this criterion were not taken into consideration for this study.
- (2) The absence of listener’s feedback cues. Only uninterrupted stretches of speech were selected for the purpose of the present analysis.<sup>3</sup> Two narratives from the total of 28 contained feedback from the interviewer, being thus discarded from the study.
- (3) (3) The quality of the recorded material. Some parts of the recordings were accompanied by extraneous noises; such material was discarded following the selection process. Only one narrative was discarded under this criterion: the interviewee accidentally hit the microphone.
- (4) (4) The length of the narrative text. Narratives had to be “short” (no longer than 5 minutes), in order to be manageable. Long narratives were not considered for the analysis. Two narratives exceeded the maximum length that was established for this study, being thus discarded from it.

Altogether, the 17 narratives in the data have a total duration of 18.5 minutes. The longest narrative, both in terms of number of words and time, has half the time that was established as the limit in the selection criteria. The shortest one has only 21.79 seconds. The average duration of the narratives in the corpus was 65.51 seconds.

It has been stated that if one wants to identify the role of prosody in the structuring of information, one must compare it with an independently obtained discourse structure, in order to minimize the risk of circularity. (SWERTS, 1997;

---

<sup>3</sup> Although narratives are the result of a jointly created process (JEFFERSON, 1978), the listener’s feedback responses, if considered in the present research, would affect the measure of speech rate values.

SWERTS; GELUYKENS, 1994). In order to have some sort of information structure against which prosody can be confronted, some authors rely on discourse segmentations resulting from discourse analysis. (GROSZ; HIRSCHBERG, 1992; GROSZ; SIDNER, 1986; PASSONNEAU; LITMAN, 1993). The problem with using the discourse analysis approach is that a priori we do not know whether it will yield more than an individual's intuition of discourse structure. If we are to depend on a discourse segmentation method, we must assure that we are employing one that is reproducible, because the more replicable a discourse segmentation model is, the stronger the evidence that discourse structure does exist.

In order to avoid the above-mentioned risk of circularity, a series of procedures was taken before the actual analysis of the data. These procedures were divided into four different stages, as described below.

The first stage covered the digitization of the acoustic material, and the transcription of the data. The narratives were digitized at 22.05 KHz with 16-bit resolution, using the speech-editing software SoundEdit 16, version 2.0 (Macromedia Inc.). They were all linearly transcribed afterwards, using standard orthography, with no punctuation marks, or special characters. Pauses were not indicated in the transcription. Incomplete words were marked with a single slash (/).

The second stage dealt with the division of the narratives into intonation units (IUs).<sup>4</sup> Five experts in Brazilian Portuguese prosody were responsible for this procedure. Each of them had access to both the transcriptions and the digital audio files of all the seventeen narratives.

The ability of labelers to agree with one another was measured using a figure called percent agreement. (GALE; CHURCH; YAROWSKY, 1992). Percent agreement averaged 92%, a result comparable to other segmentation studies. Results derived from Cochran's Q tests (COCHRAN, 1950) indicate that this agreement is significant.

The third stage tested the reliability of discourse segmentation and the applicability of the Labovian Model for this study. All the narratives, divided into intonation units, were given to seven labelers, speakers of Brazilian Portuguese, with no knowledge of discourse analysis. Participants of this experiment received an introductory text explaining the objective of the research, outlining the Labovian Model, providing a few examples and, finally, asking them to segment the narratives. It is important to note that the participants in this experiment did not have access to the audio files.

---

<sup>4</sup> An intonation unit (IU) is generally regarded as a unit composed of at least one prominent syllable with a major pitch movement. (CRUTTENDEN, 1997; CRYSTAL, 1969).



As a first step, labelers were instructed to identify the points in each narrative where the speaker had completed one communicative task. Once the segmentation was done, using the speaker's communicative intention as the criterion, they were asked to label each unit they had found according to the Labovian Model. It was assumed here that trained labelers can segment and correctly characterize narrative sections on the basis of the informational content of such sections only.

Percent agreement among labelers for the segmentation of narratives averaged 90%. Results derived from Cochran's Q tests indicate that this agreement is significant.

The agreement among labelers with regard to narrative sections was assessed with Kendall's W. The coefficient of concordance (Kendall's W) between the seven labelers was 0.73 (N=125,  $p < 0.0001$ ). This result indicates a high consistency among the subjects as a group. It should be emphasized, however, that a high or significant value of W does not mean that the agreements observed are correct. The fact that naive subjects were able to reach consensus was only taken as evidence for the reliability of the Labovian Model. Therefore, the results from this test could not be taken into account for the purpose of the main acoustic analysis.

The fourth (and last) stage corresponded to the definitive segmentation of the narratives into sections, according to Labov's Evaluative Model. Having found that the model to be used in this analysis is reliable (in the sense that it is reproducible), all narratives (transcriptions only) were then given to two experts in discourse analysis, who have previously worked with the Labovian Model, for segmentation and labeling purposes. The use of expert judgments (independent evidence) was employed as to avoid the risk of circularity. The experts were able to discuss with each other and, except for a few cases of disagreement (which were then solved by the author), they agreed with each other in more than 95% of the cases. The very few cases of inconsistencies were discussed between the experts until consensus was reached.

A number of different measurements of speech rate are found in the phonetic literature. Traditionally, speech rate has been measured as a function of words per unit of time. However, after a long debate on the inherent difficulties associated with the methods of this measurement (O'CONNELL; KOWAL, 1972), more recent research involves syllables per unit of time as the standard unit in the study of speech rate. (UHMANN, 1992; VAN DONZEL, 1999; BLAAUW, 1995). Abercrombie (1967, p.96), for example, defines speech rate as the "rate of syllable succession." This is the unit adopted by Blaauw (1995), Fon (1999), Grosjean and Deschamps (1972), Grosjean and Deschamps (1973), Grosz and Hirschberg (1992), Uhmman (1992), Goldman-Eisler (1961), Hirschberg and Grosz (1992), Van Donzel (1999), Wood (1975), to name a few. Even though, as Uhmman (1992) points out, this unit of

measurement has also the disadvantage of not taking into consideration the above-mentioned processes that are often found in rapid speech, such as assimilation and segmental deletion. Such processes may result in syllable omission, what would obviously not be covered in this unit of measurement.

The present study will opt for a measure that is mostly used in the temporal research of speech for the sake of comparability. It does recognize the pitfalls related to this choice, but assumes that they are not so serious as to invalidate the analysis. Speech rate will be interpreted here using the measurement of syllables per second.<sup>5</sup> The counting of syllables was made excluding possible contractions, as to avoid subjectivity due to perceptual factors.<sup>6</sup>

The measurement of speech rate was made by examining the waveform, on the speech-editing program Praat (BOERSMA, 2001). Pauses and nonlinguistic utterances were treated as individual units (syllables) and were included in the calculation of rate,<sup>7</sup> since, as Fon (1999, p.663) puts it, “[...] they might be indicators of conceptual planning and their existence might also contribute to rate perception.”<sup>8</sup>

## Speech rate cycle

The first step in investigating whether speech rate is used as a cue for narrative segmentation is to try to find out whether a variation in rate occurs as a function of the alternation of narrative sections. It was verified in Oliveira Júnior (2002a) that a cycle of varying fluency (stated as a measure of pause to speech ratio) occurs in spontaneous narratives, and that this cycle not only reflects the cognitive process of planning and execution, as proposed by Henderson, Goldman-Eisler, Skarbek (1966), Goldman-Eisler (1967), and Butterworth e Goldman-Eisler (1979), but also emerges as a function of the way narratives are structured.<sup>9</sup> Since pause occurrence and duration often have a decisive influence on speech rate, it is expected that a “speech rate cycle” also emerge in spontaneous narratives.

---

<sup>5</sup> For a detailed discussion on various units of speech rate measurement and on the difference between speech rate and articulation rate, see Oliveira (2000).

<sup>6</sup> For a discussion on how subjectivity due to perception of contractions may interfere on measurements of speech rate, see Uhmman (1992, p.312).

<sup>7</sup> Note that the pauses that occur at the end of the intonation units were considered to be part of them.

<sup>8</sup> An alternate calculation that did not include pauses and nonlinguistic utterances was also undertaken. The results obtained from both measurements revealed that the difference was not significant. The measurement including pauses and nonlinguistic utterances was adopted for methodological reasons. (FON, 1999).

<sup>9</sup> Fitting the Labovian model of narrative analysis.

In a study on speech rate as a reflection of variance and invariance in conceptual planning in storytelling, Fon (1999, p.666) analyzed the elicited narratives of ten speakers of Mandarin and found that they were generally sensitive to different story structures and that, as a rule, they accommodated their speech rate as to reflect these structures. She concluded that “[...] invariance of speed lies in the fluctuating patterns and its correlation with story parts.” The narratives analyzed by Fon (1999) were, like many other correlated studies, elicited from cartoons. A total of two sets of four-frame cartoon strips were presented to the subjects: one displaying an AAAB structure and the other displaying an ABCD structure.<sup>10</sup> In terms of story segmentation, it was verified a clear one-to-one correlation between frames and story parts in the narration of the AAAB-type cartoon; the ABCD-type cartoon, on the other hand, did not display such a correspondence. As for the correlation of story part and rate cycle, it was observed that a story part can be either subsumed with a rate cycle, or it can span across two or more cycles.

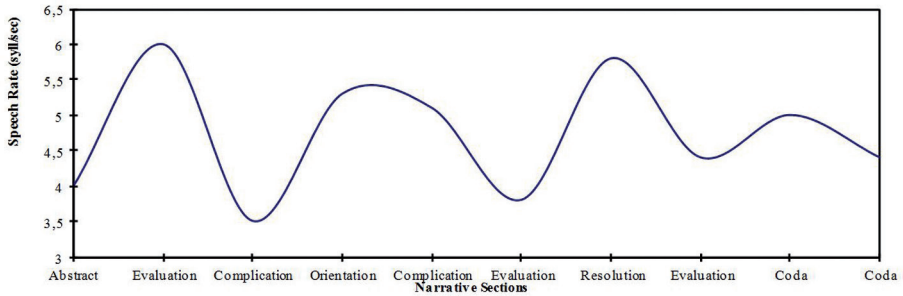
The type of narrative analyzed by Fon (1999) is obviously different from the type of narrative utilized in the present investigation, so the fact that a cycle of varying rate reflecting the narrative structure was found in her study does not necessarily imply that the same will be verified here. In Fon’s study, subjects were constrained by fixed sets of cartoon frames: the boundaries of story parts were visually indicated in the eliciting material. Thus, despite the fact that in one of the cartoon types a clear correlation between frames and story part was not verified, the narratives in her data might as well reflect the structure that was visually imposed by the comic strips. Consequently, it may be argued that the occurrence of a speech rate cycle in this particular case, rather than indicating the awareness of narrative structure by the tellers, actually reflects the graphical characterization of story parts in the cartoons. If this cycle of varying rate is reproduced in the present data, which is composed by non-elicited, spontaneous narratives, corroboration for Fon’s findings will be provided more convincingly.

A typical rate cycle in the narratives of this investigation is given in Figure 1 below:

---

<sup>10</sup> The AAAB-type cartoon had five story parts (P): couple getting married (P1), first sign holder (P2), second sign holder (P3), third sign holder (P4), pastor’s word (P5). Both P1 and P2 pertain to the first frame. All the other descriptions have a one-to-one correlation with the frames: P3 with Frame 2, P4 with Frame 3, and P5 with Frame 4. The ABCD-type cartoon had four story parts: the clinic (P1), no business (P2), change sign (P3), husband pulling wives/good business (P4). P1 and P2 refer to the first frame. P3 refers to both the second and the third whereas P4 refers to the last frame. Unlike the AAAB-type cartoon, there is no clear one-to-one correlation between frames and story parts in the ABCD-type cartoon.

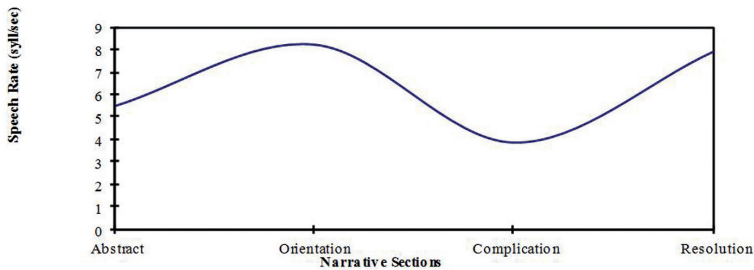
**Figure 1 – Speech rate plotted over time for narrative 14**



**Font:** Prepared by the author.

The fluctuation of speech rate in this narrative, by a function of its integrant sections, is quite evident. There seems to exist a tendency in storytelling to segment sections by means of manipulating speech rate. In some cases, a clear pattern of slow-fast speech occurs, as in Narrative 15 (Figure 2 below):

**Figure 2 – Speech rate plotted over time for narrative 15**

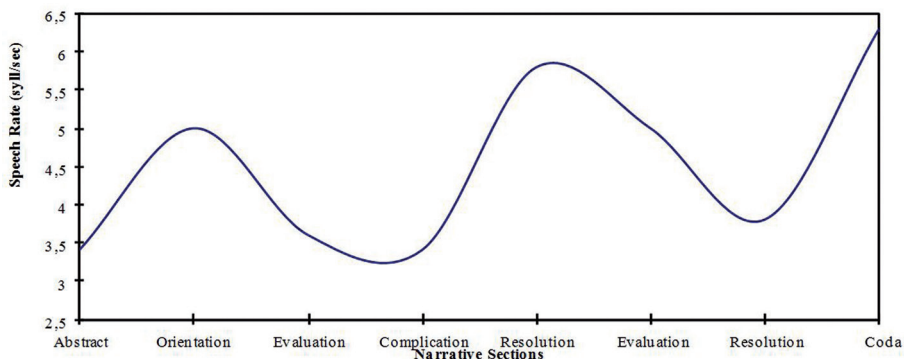


**Font:** Prepared by the author.

Cases exhibiting such a precise polarity distinguishing narrative sections are very rare, though. Generally, a pattern displaying at least one section that does not differ from the preceding one by means of a diametric relation is much more common. The evaluation section that comes after the second complication section in narrative 14 (Figure 1), for example, is not characterized by a rate of diametric value, but rather by a value that has a symmetrical relation with it. Therefore, instead of a fast-slow-fast-slow cycle representative of narrative 15 (Figure 2), that specific point in narrative 14 is characterized by a fast-slow-slower-fast cycle. Still, a difference in rate is easily verified, but this difference is not an asymmetric one. It should be pointed out that this does not seem to be a haphazard phenomenon: sections that follow the upward or downward movement of the previous section

in terms of speech rate value are characterized by their evaluative content. Since evaluative sections (or any other section that presents a high amount of evaluative features, for that matter) present in most part a high degree of embedment in other sections, it seems reasonable to expect that they follow the upward or downward direction of the section in which they are embedded. Narrative 17 (Figure 3) reiterates this point:

**Figure 3 – Speech rate plotted over time for narrative 17**



**Font:** Prepared by the author.

The two sections that do not present an asymmetric relation with the previous ones are both evaluations. This example reflects quite accurately what can be found in most narratives in the present data.

Therefore, from a global perspective, narrative structure seems to be manifested by means of variation in speech rate, as demonstrated above. Speakers apparently indicate a change in narrative section by shifting the rate of their speech. This maneuver results in a cycle similar to the one proposed by Henderson, Goldman-Eisler and Skarbek (1966), Goldman-Eisler (1967), and Butterworth and Goldman-Eisler (1979) for the variation in pause to speech ratio. If the “cognitive cycle” predicts that speech is more hesitant as a result of the cognitive process of planning and more fluent as a consequence of the execution of the plan made in the hesitant phase, then it should be reasonable to expect that the variation in speech rate reflect the fact that speech is slower when concepts are being formed and faster when the concepts are being verbalized.

The “speech rate cycle” found in the present investigation corroborates Fon’s (1999) hypothesis that speech rate reflects how conceptual planning is laid out during speech. If statistically significant differences in speech rate among the various sections in the narratives are found, this hypothesis will be further substantiated. However, before the investigation of whether speech rate can be

used as a tool for indicating conceptual coherence is accomplished, differences in speech rate values at the local level will be studied. The question here is whether intonation unit boundaries that correspond to narrative boundaries present a higher difference in rate than intonation unit boundaries that do not function as narrative boundaries. If a higher amount of rate difference is found in the local level, the hypothesis that speech rate is used as a cue for narratives segmentation will be then ratified.

### **Speech rate reset**

Previous works demonstrated that the occurrence of pauses and their duration can predict quite accurately the presence of a narrative boundary on the local level: pauses tend to be longer than average when they occur at an IU boundary that separates two narrative sections. (OLIVEIRA JÚNIOR, 2002b). It would be interesting to investigate whether speech rate has also a decisive role in the characterization of narrative boundaries. In order to find out if this is the case, a new unit of measurement will be introduced here: the rate difference, or "rate reset," which can be defined as the distance in terms of syllables per second between the speech rate values before and after an intonation unit boundary. The assumption to be tested then is whether breaks between narrative sections can be signaled by means of rate discontinuity. Based on the high correlation between speech rate and pausal phenomena, it is expected that speech rate reset will be higher at narrative boundaries than elsewhere in a narrative text.

Rate reset was computed as the difference between the speech rates of two adjacent IUs. Only the absolute values are taken into consideration for the purpose of the statistical analysis. Results from a t-test showed that rate reset values do not differ significantly for the narrative and the non-narrative boundaries ( $t=0.255$ ,  $df=620$ ,  $p=0.7986$ ).

Therefore, although there exists a high correlation between longer pauses and higher speech rate reset, the employment of the former as an indication of narrative boundary does not necessarily mean the occurrence of the latter. Speech rate is only used as a segmentation tool at the global level. This can be verified by the employment of a fluctuation pattern of varying rate values that form a cycle corresponding, in most cases, to the way narratives are structured into semantically individualized sections. Speech rate is not employed at the local level as a cue for narrative segmentation.

The following section will explore the possibility of relating speech rate to individual narrative sections.

## Speech rate as a representation of narrative section

In an investigation of forms and functions of speech rate in conversation, Uhmman (1992) suggests that participants make systematic use of changes in speech rate in order to contextualize their utterances in a certain way. According to her study, speech rate aids in the semantic task of information structure by distinguishing highly relevant parts in a talk from less central or less relevant parts. She found, for example, that fast speech (in terms of syllables per second) serves to contextualize parenthesis, side-sequences, repair sequences, afterthoughts as turn-exit devices, and parts of minor relevance for the development of the speaker's argument; slow speech, on the other hand, characterizes parts of major relevance in speech.

Obviously the criteria that are used to establish what is relevant and what is not can vary greatly, mainly because this distinction, rather than being a dichotomic one, actually reflects a scalar notion that is directly associated to a certain context. In her study for conversation, Uhmman (1992, p.326) proposes that the notion of relevance is closely related to topicality:

a turn is more relevant if it contains a contribution to the ongoing topic that is not already known to the recipient due to one or more of the following reasons: (a) it was already mentioned in the prior discourse, (b) it summarizes prior arguments, and (c) it gives some sort of information which already belongs to the recipient's knowledge for other reasons.

It seems, then, that Uhmman's working assumption for relevance is connected with the well-known distinction of given-new information.

In narratives, the concept of relevance could be straightforwardly associated with the role that each individual section plays in the story. The Labovian complicating action, which brings a description of the most important events in the narrative, could then be regarded as relevant information and, according to the hypothesis discussed above, would present a relatively slower rate than sections such as orientation, abstract and codas, which are for the most part characterized by propositions that elaborate the events described in the complicating action. Codas, abstracts and orientations would present a faster rate, according to what is hypothesized above. Resolutions, on the other hand, are composed of narrative clauses, and thus would present a rate similar to the complicating action. The status of evaluative sections, however, is somewhat dubious. If one considers evaluations as propositions that are outside the narrative sequence, serving as background information that is not necessarily pertinent to the comprehension of the story as a whole, then such sections could be regarded as not relevant, according to the notion of relevance discussed

above, and would, for that reason, be grouped with the abstract, orientation and coda. Conversely, if evaluations are viewed as the “raison d’être” of a narrative – as Labov (1972) defines them, they could be then grouped with the complicating action and the resolution, forming a group of the most important (or relevant) information in a narrative. Since the present analysis takes the Labovian model as the conceptual working frame, evaluations will be grouped with complicating actions and resolutions. The assumption then is that evaluative sections will present a slower speech rate.

Table 1 below provides the speech rate mean values for each narrative section in the data.

**Table 1 – Mean speech rate values and standard deviations for each of the six narrative sections (in syllables per second)**

	<i>Abstract</i>	<i>Orientation</i>	<i>Complication</i>	<i>Evaluation</i>	<i>Resolution</i>	<i>Coda</i>
<i>Mean</i>	6.2	5.6	5.5	5.5	5.5	6.1
<i>Sd.</i>	2.2	1.1	1.2	1.8	1.8	1.6

**Font:** Prepared by the author.

Although the differences among sections are not statistically significant ( $F(5,91)=0.524, p=0.7573$ ), a trend emerges. Complications, evaluations, and resolutions form a group of similar lower values; abstracts and codas form another group of relatively higher values. Orientations, as opposed to what was expected, are in general characterized by a slower rate. They are grouped with the sections that, according to the concept of relevance discussed above, are more relevant in a narrative text.

A closer look at the narratives that display lower speech rate values in the orientation section reveals that in most cases the information that is conveyed in such sections are fragmented and present all manner of hesitation phenomena, such as long pauses, repairs, false starts, etc. These facts seem to contribute to the slower rate of speech in the orientations. In Example 1 below, for instance, the orientation section of the narrative is uttered in a much slower rate in comparison to the other sections. This section is, however, characterized by fragmented information, a false start and by the incidence of longer pauses at the end of almost all IUs:



**Example 1 – Excerpt of Narrative 08. The narrative is segmented in terms of IUs (numbered on the right) and sections (separate rows correspond to separate sections). Numbers in parenthesis correspond to pause duration (in ms.)**

<b>Narrative 08</b>
<i>Orientation</i>
02 - tava passando no parque ali (0.85)
03 - foi dois mo/ dois moleques lá
04 - roubaram tudo dele (0.85)
05 - tava voltando do inglês (0.62)
06 - chegou em casa chorando (0.28)
<i>Orientation</i>
07 - daí tava eu e meu outro irmão em casa (0.99)

<b>Narrative 08</b>
<i>Orientation</i>
02 - <i>he was passing by the park over there (0.85)</i>
03 - <i>two bo/ two boys there</i>
04 - <i>they stole everything from him (0.85)</i>
05 - <i>he was coming back from English class (0.62)</i>
06 - <i>he arrived at home crying (0.28)</i>
<i>Orientation</i>
07 - <i>so there we were me and my other brother at home (0.99)</i>

**Font:** Prepared by the author.

The higher speech rate values of codas and abstracts, on the other hand, might indicate the existence of a possible “narrative frame,” marked by means of acceleration in speech rate. In their studies on sequential temporal patterns in elicited narratives, Henderson, Goldman-Eisler and Skarbek (1966) noted that in both the spontaneous and the read-aloud versions of their narratives, a period of “rambling introductions and tailpieces” could be easily verified. According to their study, the read-aloud narratives present these “entry and exit phenomena” because of the cognitive act of “scanning ahead” associated with reading. They do not offer an explanation for the occurrence of such phenomena in the spontaneous versions of the narratives as well. Brubaker (1972), on the other hand, found a statistically significant effect for speech rate in relation to sentential position only at the end of reading passages. According to him, “subjects tended to speed up in their performance as they neared the end of the passage, presumably in order to terminate the laboratory task more quickly.”<sup>11</sup>

It is suggested here that sections displaying a higher rate surround spontaneous, non-elicited narratives as a way of indicating the limits of this type

<sup>11</sup> See also Uhmman (1992), who found the rate of afterthoughts and summaries as topic-exit device to be faster than average in her study.

of discourse that is monological in nature: it is well documented that narrative texts not only require an extended turn in a conversation but also the suspension of turn exchanging. (SCHIFFRIN, 1994).

The acceleration of speech rate that occurs at the beginning of a narrative might be a cue for the listener that the turn that is about to begin is a possibly long one and that its non-interruption would be desired. In a conversation, the speeding up at the beginning of a narrative may also be interpreted as a technique of “grabbing the conversational turn”. (SELTING, 1996).

The high rate at the end of a narrative, on the other hand, is much more related to the content conveyed in coda sections. The coda signals the “sealing off” of a story by revealing the effects of the events on the narrator. It is used as a device to reinstate the conversational mode and is often characterized by the communication of information that is not directly relevant to the events reported in the story. (LABOV; WALETZKY, 1967, LABOV, 1972). As previously pointed out, non-relevant information is regularly uttered in a faster rate, which would justify the speeding up in coda sections.

Furthermore, the fact that evaluations are often uttered at a slow rate corroborates the assumption that they carry relevant information in a narrative.<sup>12</sup>

Note that the comments made thus far concerning the relation between speech rate and narrative sections should be regarded much more as speculation than observation on factual phenomena. Recall that no statistically significant effect was found to corroborate the existence of such a connection. The numbers only suggest that a trend on that direction may be present. A larger amount of data would be necessary in order to validate the premises that were discussed above. Of course that this is not to say that there does not exist a connection between speech rate and information in discourse, but that in narratives, such association could not be statistically verified on a more global level. The examination of information on a local level may result in a more clear understanding of the relationship between speech rate and information in discourse. So, for example, if pieces of information that are included in a narrative section were taken separately and their rate values were considered under a discourse analytic perspective (using, for example, the independent model of discourse structure developed by Grosz and Sidner (1986) or the ‘Information Structure In Discourse – ISID’ model, proposed by Van Donzel (1999), the results of the statistical analysis could differ greatly from the ones in the present investigation. Although such analysis is not the goal of this study,

---

<sup>12</sup> See, however, Koopmans-van Beinum and Van Donzel (1996), who found low average syllable duration (ASD) values connected with “expansions in the form of personal comments of the speaker on the manner of retelling the story (e.g., ‘I don’t remember that exactly’), or comments on the whole situation (e.g., ‘just as people can do in such a situation’).” Low ASD values correspond to faster rate.

a few examples will be given in order to illustrate that in many cases a clear correspondence between content and speech rate can be established.

The example to follow was extracted from the orientation section of narrative 05. The teller was trying to remember exactly when the events he is reporting took place:

**Example 2 – Excerpt of Narrative 05. The narrative is segmented in terms of IUs (numbered on the right) and section. Numbers in parenthesis correspond to pause duration (in ms.). Numbers in the SR column correspond to the rate in syllables per second for each IU**

Narrative 05	SR	Narrative 05	SR
<i>Orientation</i>		<i>Orientation</i>	
18 - aí acho que em noventa e/	3.7	18 - <i>then I think that in ninety/</i>	3.7
19 - eu tinha catorze anos	10.4	19 - <i>I was fourteen</i>	10.4
20 - isso devia ser éh oitenta e quatro	7.7	20 - <i>that must've been uh eighty four</i>	7.7
21 - oitenta e cinco	7.4	21 - <i>eighty five</i>	7.4
22 - por aí (0.28)	3.3	22 - <i>some time around this (0.28)</i>	3.3

**Font:** Prepared by the author.

Orientation sections are in general characterized by a lower speech rate in narratives, as previously discussed. In this narrative, the orientation section is uttered in a rate below the average value for the whole story (4.8, as opposed to 5 syllables per second), being only faster than the Resolution section (which is uttered in a rate of 4.1 syllables per second). However, IU 19, which is located in the middle of the orientation section, has the highest rate value in the narrative. The reason for this is only clear if one takes into account that IU 19 is actually a self-repair.<sup>13</sup> Self-repairs are commonly uttered in a faster rate for contextual reasons: the speaker wants to indicate that the space occupied by a self-repair in the conversation is as small as possible and will not compromise her or his turn as a whole. (UHMANN, 1992). Observe that IUs 20 and 21 also present a rate faster than the average. Since both of them serve as rectifications of the information given in IU 18, they can be also typified as self-repairs. The rate in 22 drops considerably, marking the end of the repair and the return to the narration.

Faster speech rate is also employed when the speaker is making a parenthetical remark, or side comment, during the course of the story. Since both parentheses

<sup>13</sup> The self-repair in this case is signaled by means of a false start (see SACKS; SCHEGLOFF; JEFFERSON, 1977 for a discussion on the various forms of introducing self-repairs in conversation).

and interpolated information interrupt the narrative, they are often marked as dissimilar from the adjacent passages. This is mainly achieved by means of variation in prosody. Speech rate seems to be one of the most effective strategies employed for this purpose. (UHMANN, 1992). Some examples of variation on speech rate as a result of the occurrence of parentheses and side comments are given below:

**Example 3 – Excerpt of Narratives 09, 16 and 01. The narratives are segmented in terms of IUs (numbered on the right) and sections (separate rows correspond to separate sections). Numbers in parenthesis correspond to pause duration (in ms.). Numbers in the SR column correspond to the rate in syllables per second for each IU**

<b>Narrative 09</b>	SR
<i>Complication</i>	
08 – e éh aí eu acordei	4.2
09 – tava dormindo	7.5
10 – acordei (1.22)	1.7

<b>Narrative 09</b>	SR
<i>Complication</i>	
08 – and uh then I woke up	4.2
09 – I was sleeping	7.5
10 – I woke up (1.22)	1.7

<b>Narrative 16</b>	SR
<i>Evaluation</i>	
26 – porque... ele queria/	5.5
27 – lógico que arrumou outra perua	7.3
28 – ir com né? com ela na festa	5.7

<b>Narrative 16</b>	SR
<i>Evaluation</i>	
26 – because... he wanted/	5.5
27 – obviously he found another bitch	7.3
28 – to go you know? to the party with her	5.7

<b>Narrative 01</b>	SR
<i>Orientation</i>	
02 – ele tava passando por uma rua um/ (0.46)	5
03 – a gente chama de de portão do gelo (0.40)	5.5
04 – esse lugar (0.63)	3.5
<i>Evaluation</i>	
05 – e agora ele tá bem mudado	7.8
06 – começa/ fizeram muita casa	8.6
07 – mas ag/ antes (0.64)	2.5
08 – ele era muito cheio de de árvore	5.1

<b>Narrative 01</b>	SR
<i>Orientation</i>	
02 – he was passing by a street (0.46)	5
03 – we call “the gate of the ice” (0.40)	5.5
04 – that place (0.63)	3.5
<i>Evaluation</i>	
05 – and now it’s pretty changed	7.8
06 – they began/ they built many houses	8.6
07 – but now/ before (0.64)	2.5
08 – it was full of trees	5.1

<b>Narrative 01</b>	SR
09 – de de mato	5.2
10 – não de árvore	7.7
11 – de mato	7.5
12 – era matagal (0.73)	3.5
<i>Complication</i>	
13 – e e pai vinha por ali de noite (0.98)	3.1

<b>Narrative 01</b>	SR
09 – <i>of grass</i>	5.2
10 – <i>not trees</i>	7.7
11 – <i>grass</i>	7.5
12 – <i>it was a place full of grass (0.73)</i>	3.5
<i>Complication</i>	
13 – <i>and my father was coming along in that night (0.98)</i>	3.1

**Font:** Prepared by the author.

IU 09, in the excerpt of narrative 09 given in Example 3 above, obviates what can be inferred from the information given in IU 08. It is a parenthesis because it discontinues the flow of the events, but at the same time it constitutes redundant material. The fast rate is a direct result of the status that this IU occupies in the narrative as a piece of superfluous information. The excerpt of narrative 16 (also in Example 3 above) is similar to the one extracted from narrative 09 only because the fastest IU communicates something that should have been mentioned previously, but was not. A parenthesis was necessary in this case in order to make the argument understandable. Therefore, it is not solely the importance of the information that dictates the rate of speech, but also the status of the information on a discursive level.

The last excerpt in Example 03, that of narrative 01, is a very interesting instance in which an entire section functions as a side comment. The section is actually an external evaluation and, instead of being uttered in a slow mode, following the general trend of evaluative sections, as discussed above, this particular section has a speech rate value higher than the ones surrounding it. This is probably because it interrupts the narrative, deviating the listener's attention to the setting of the story, rather than to the actions.<sup>14</sup> Since the actions are obviously the most important element in a narrative, the information given in side comments are to be interpreted as nonessential, which result in their being uttered in a fast mode.

It seems, then, that the information conveyed at the local level is of much more importance for the determination of speech rate than that gathered in a more global discursive level. This could explain why, in some cases, the rate

---

<sup>14</sup> Note that since this section provides information about the place where the events took place, it could be easily classified as an orientation. However, because its primary function is, rather than to provide the necessary background information of the setting where the events took place, to enhance the point of the narrative (by creating a creepy atmosphere), it was considered to be essentially evaluative.

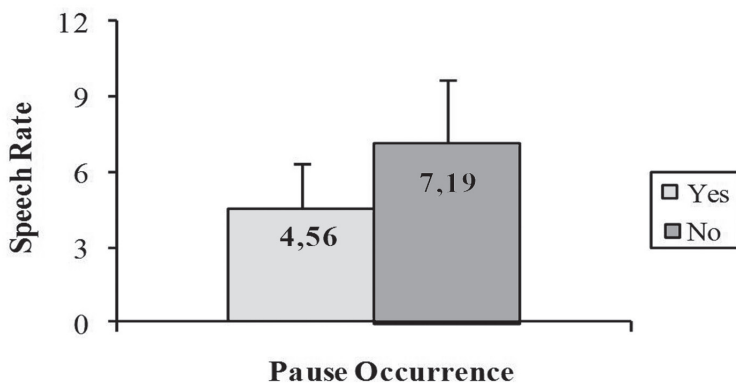
of a given section does not follow the trend that was verified for the narratives in the data, a trend that for the most part agrees with the concept of relevance discussed above. It is not the primary function of a narrative section that always determines the rate of the section: the elements within the section should be taken into consideration all the time.

### Correlation of pause and speech rate

If it is true that one of the major determinants of speech rate is pause occurrence and pause duration, as the literature suggests (confer GOLDMAN-EISLER, 1956; GOLDMAN-EISLER, 1968; GROSJEAN, 1980; SABIN, 1976; SABIN, et al. 1979), it is to be expected that both phenomena are strongly correlated with speech rate.

The first hypothesis to be tested is whether pause occurrence determines the value of speech rate. The assumption is that the occurrence of a pause at the end of an IU will trigger a lower speech rate. Figure 3.4 below brings together the mean values of speech rate in the presence and absence of pause:

**Figure 4 – Speech rate as a function of pause occurrence**



**Font:** Prepared by the author.

Speech rate tends to be higher in the presence of pause, and much lower when pause does not occur. The difference between the two conditions is significant ( $t=14.677$ ,  $df=606$ ,  $p<0.0001$ ).

As for pause duration, the hypothesis follows that there is an inverse relationship between pause duration and speech rate: the longer the total pause duration within an IU, the lower the speech rate value for that IU will be. However,

the correlation between the values of speech rate and pause duration per IU was not very high for this data ( $r=-0.52$ ,  $N=627$ ,  $p<0.0001$ ).

## **General discussion and conclusions**

The present paper investigated the role of speech rate phenomena in narrative texts, focusing primarily on how the temporal dimension of speech helps in the characterization of narrative structure. The following research questions were put forward:

- (1) Are speech rate phenomena systematically manipulated in storytelling in order to make the structure of narrative texts more transparent?
- (2) If so, in exactly what way is narrative structure reflected by means of manipulation in speech rate?
- (3) Are different narrative sections characterized by particular speech rates?

In order to answer the first two questions, analyses on both the global and local levels of the narratives were carried out. Both analyses were working on the assumption that, if speech rate was used as a cue for narrative segmentation the same way pausal phenomena are utilized (OLIVEIRA JR., 2002b), this would be reflected on at least one of these levels.

Based on the findings by Fon (1999), and on what was verified in a previous work with regard to the variation in pause to speech ratio in narratives, the “cognitive rhythm” (OLIVEIRA JÚNIOR, 2002a), it was hypothesized that a cycle of varying rate would also be present in the narratives, reflecting the way they are structured. Variation in rate was examined by taking into consideration the Labovian model of narrative analysis. The results indicated that speech rate values fluctuate considerably at the global level, resulting in a cycle very similar to the one observed for the pause to speech ratio.

The “rate cycle” is not itself a new finding. Fon (1999), for example, has demonstrated that it occurs quite regularly in elicited narratives, reflecting a correlation between cycles of varying rate and story parts, which can either span across cycles or be subsumed within one cycle. However, no attempt has been made so far to relate this observable phenomenon to narrative structure, using spontaneous, non-elicited narratives as the empirical database. By taking an independent model of narrative analysis into account and trying to connect it with speech rate variation, it was demonstrated that there exists a one-to-one correlation between narrative sections and rate cycle. This finding strengthens the importance of the temporal prosodic phenomena in the segmentation of narrative texts

Rate, on the other hand, did not prove to be a reliable tool for the signaling of narrative section boundaries on the local level. It was hypothesized that the difference in rate between two intonation units that coincided with a narrative boundary would be greater than elsewhere. So, for example, it was expected that a storyteller uttered the last IU of a narrative section in a way that would differ quite noticeably from the first IU of the following section, so as to indicate a change of sections by means of speech rate reset. This feature would serve, along with pause duration, as a cue to narrative segmentation. Statistical analyses, however, showed no significant effect for speech rate reset as a narrative section boundary marker.

It was also hypothesized that speech rate varies as a function of the message conveyed at the global level within narrative sections. Speech rate has been often related to levels of relevance in textual analyses (UHMANN, 1992): the faster someone speaks, the less relevant the content of what is being uttered, and vice-versa. Based on this assumption, it was expected that a close relationship between slower rates and crucial narrative sections (the complicating action and the evaluation) would be found. Although a trend in this direction could be verified, statistical analyses revealed that differences in rate between narrative sections are not significant.

Finally, a correlation was made between speech rate and pause occurrence/duration in the data. The results of these correlations suggest that the occurrence of a pause within an IU guarantees that the speech rate of that IU will be higher than it would be if no pause was employed. On the other hand, the duration of the pause does not seem to have a straightforward relation with the rate of an IU. Contrary to what was expected, the correlation between pause duration and speech rate in an IU was not found to be significant.

OLIVEIRA JÚNIOR, M. Um estudo sobre a velocidade de fala como marca de segmentação em narrativas espontâneas. *Alfa*, São Paulo, v.56, n.2, p.623-651, 2012.

- *RESUMO: A velocidade da fala é analisada neste trabalho como um recurso prosódico empregado na sinalização da estrutura da narrativa espontânea. Partindo do pressuposto de que a narrativa tem uma estrutura interna, e de que os pares envolvidos em interações guiam seus movimentos conversacionais e suas atividades mais globais, com o propósito de torná-la evidente (JEFFERSON, 1978; SACKS, 1972), o presente trabalho examina fenômenos pertinentes à velocidade de fala, numa abordagem acústico-experimental, em 17 narrativas espontâneas, usando para isso um dos modelos mais influentes na análise da narrativa - o modelo laboviano. (LABOV, 1972). A variável prosódica sob investigação é analisada em dois níveis diferentes: em pontos específicos nas narrativas correspondentes aos limites de seções narrativas (nível local), e dentro de diferentes seções nas narrativas como um todo (nível global). Os resultados indicam que a velocidade de fala atua exclusivamente no nível global, gerando um padrão cíclico de diferentes velocidades ao longo do texto narrativo.*



*A velocidade de fala, no entanto, não caracteriza seções narrativas e não é manipulada no nível local, indicando limites de seções narrativas.*

- PALAVRAS-CHAVE: *Narrativa espontânea. Prosódia. Estrutura do discurso. Velocidade de fala.*

## REFERENCES

ABERCROMBIE, D. *Elements of general phonetics*. Edinburgh: Edinburgh University Press, 1967.

ARCHAKIS, A.; PAPAZACHARIOU, D. Prosodic cues of identity construction: intensity in Greek young women's conversational narratives. *Journal of Sociolinguistics*, London, v.5, n.12, p.627-647, 2008.

ATTARDO, S.; PICKERING, L. Timing in the performance of jokes. *Humor*, Berlin, v.2, n.24, p.233-250, 2011.

BARTHES, R. An introduction to the structural analysis of narrative. *New Literary History*, Baltimore, n.6, p.237-230, 1975.

BLAAUW, E. *On the perceptual classification of spontaneous and read speech*. 1995. 224f. (Dissertation) - Utrecht University, Utrecht, 1995.

BLACK, J.; WLILENSKY, R. An evaluation of story grammars. *Cognitive Science*, Norwood, n.3, p.213-230, 1979.

BOERSMA, P. Praat, a system for doing phonetics by computer. *Glott International*, London, n.5, p.341-345, 2001.

BROWN, G.; CURRIE, K.; KENWORTHY, J. *Questions of intonation*. London: Croom Helm, 1980.

BRUBAKER, R. S. Rate and pause characteristics of oral reading. *Journal of Psycholinguistic Research*, New York, v.2, n.1, p.141-147, 1972.

BUTTERWORTH, B.; GOLDMAN-EISLER, F. Recent studies on cognitive rhythm. In: SIEGMAN, A.; FELDSTEIN, S. *Of Speech and Time*. Hillsdale: Lawrence Erlbaum, 1979.

COCHRAN, W. G. The comparison of percentages in matched samples. *Biometrika*, London, n.37, p.256-266, 1950.

COLLIER, R.; PIYPER, J. R. D.; SANDERMAN, A. Perceived prosodic boundaries and their phonetic correlates. In: ARPA WORKSHOP ON HUMAN LANGUAGE TECHNOLOGY, 1993. *Proceedings...* Plainsboro, 1993. p.341-345.

CRUTTENDEN, A. *Intonation*. Cambridge: Cambridge University Press, 1997.

CRYSTAL, D. *Prosodic systems and intonation in english*. Cambridge: Cambridge University Press, 1969.

FERRÉ G. Gesture catchments and density in narratives of personal experience. In: GESPIN: GESTURE AND SPEECH IN INTERACTION, 2009, Poland. *Proceedings...* Poland, 2009. p.1-7

FON, J. Speech rate as a reflection of variance and invariance in conceptual planning in storytelling. In: THE INTERNATIONAL CONGRESS OF PHONETICS SCIENCES, 14., 1999. *Proceedings...* San Francisco, 1999. p.663-666.

FUJISAKI, H. Prosody, models, and spontaneous speech. In: SAGISAKA, Y; CAMPBELL, N.; HIGUCHI, N. *Computing prosody: computational models for processing spontaneous speech*. New York: Springer, 1997. p.27-42.

GALE, W.; CHURCH, K. W; YAROWSKY, D. Estimating upper and lower bounds on the performance of word-sense disambiguation programs. In: THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 30., 1992. *Proceedings...* Newark, 1992. p.249-256.

GENETTE, G. *Narrative Discourse*. Ithaca: Cornell University Press, 1980.

GLENN, C. G. The role of episodic structure and of story length on children's recall of single stories. *Journal of verbal learning and verbal Behavior*, New York, n.17, p.229-247, 1978.

GOLDMAN-EISLER, F. *Psycholinguistics: experiments in spontaneous speech*. London; New York: Academic Press, 1968.

\_\_\_\_\_. Sequential temporal patterns and cognitive processes in speech. *Language and Speech*, Middlesex, v.2, n.10, p.122-132, 1967.

\_\_\_\_\_. The rate of changes in the rate of articulation. *Language and Speech*, Middlesex, n.4, p.171-174, 1961.

\_\_\_\_\_. The determinants of the rate of speech output and their mutual relations. *Journal of the Psychosomatic Research*, Oxford, n.1, p.137-143, 1956.

GROSJEAN, F. Temporal variables within and between languages. In: DECHERT, H.; RAUPACH, M. *Towards a cross-linguistic assessment of speech production*. Frankfurt: Lang, 1980. p.39-53.

GROSJEAN, F.; DESCHAMPS, A. Analyse des variables temporelles du français spontané. II. Comparaison du français oral dans la description avec l'anglais (description) et avec le français (interview radiophonique). *Phonetica*, Basel, n.28, p.191-226, 1973.

\_\_\_\_\_. Analyse des variables temporelles du français spontané. *Phonetica*, Basel, n.26, p.126-156, 1972.

GROSZ, B.; HIRSCHBERG, J. Some intonational characteristics of discourse structure. In: THE INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 1992. *Proceedings...* Banff, 1993. p.429-432.

GROSZ, B.; SIDNER, C. L. Attentions, intentions, and the structure of discourse. *Computational Linguistic*, Arlington, n.85, p.363-394, 1986.

HENDERSON, A.; GOLDMAN-EISLER, F.; SKARBEEK, A. Sequential temporal patterns in spontaneous speech. *Language and Speech*, Middlesex, n.9, p.207-216, 1966.

HIRSCHBERG, J.; GROSZ, B. Intonation features of local and global discourse structure. In: DARPA WORKSHOP ON SPOKEN LANGUAGE SYSTEMS, 1992. *Proceedings...* Arden House, 1992. p.441-446.

HWANG, C. H.; SCHUBERT, L. K. Tense trees as the 'fine structure' of discourse. In: THE ANNUAL MEETING OF THE ACL, 30., 1992. *Proceedings...* Newark, 1992. p.232-240.

JEFFERSON, G. Sequential aspects of storytelling in conversation. In: SCHENKEIN, J. *Studies in the organization of conversational interaction*. New York: Academic Press, 1978. p.219-248.

KOOPMANS-VAN BEINUM, F. J.; VAN DONZEL, M. E. Discourse structure and its influence on local speech rate. In: THE INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 1996. *Proceedings...* Philadelphia, 1996. p.1724-1727.

LABOV, W. The transformation of experience in narrative syntax. In: \_\_\_\_\_. *Language in the inner City*. Philadelphia: University of Pennsylvania Press, 1972. p.354-98.

\_\_\_\_\_. The study of language in its social context. In: FISHMAN, J. *Advances in the Sociology of Language*. The Hague: Mouton, 1971. p.152-216.

LABOV, W.; WALETZKY, J. Narrative analysis: oral versions of personal experience. In: HELMS, J. *Essays on the verbal and visual arts*. Seattle: University of Washington Press, 1967.

LAAN, G. P. M. The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a reading speaking style. *Speech Communication*, Amsterdam, v.22, p.43-65, 1997.

LEECH, G.; SHORT, M. *Style in fiction: a linguistic introduction to english fictional prose*. London: Longman, 1981.

LEHISTE, I. Some phonetic characteristics of discourse. *Studia Linguistica*, Lund, n.36, v.2, 1982. p.117-130.

\_\_\_\_\_. The phonetic structure of paragraphs. In: COHEN, A.; NOOTEBOOM, S. *Structure and process in speech perception*. Berlin: Springer-Verlag, 1975. p.195-206.

LEVIN, H.; SCHAFFER, C.; SNOW, C. The prosodic and paralinguistic features of reading and telling stories. *Language and speech*, Middlesex, v.25, n.1, p.43-54, 1982.

MANDLER, J. M.; JOHNSON, N. S. Remembrance of things parsed: story structure and recall. *Cognitive Psychology*, New York, n.9, p.111-151, 1977.

O'CONNELL, D. C.; KOWAL, T. D. Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research*, New York, n.1, p.155-164, 1972.

O' CONNELL, D. C.; KOWAL, S. Pausology. In: SEDELOW, W.; SEDELOW, S. *Computers in Language Research*. Amsterdam: Mouton Publishers, 1983. p.221-301.

OLIVEIRA JÚNIOR., M. Pausing strategies as means of information processing in spontaneous narratives. In: INTERNATIONAL CONFERENCE ON SPEECH PROSODY, 1., 2002. *Proceedings...*, Aix-en-Provence, 2002a. p.539-542.

\_\_\_\_\_. The role of pause occurrence and pause duration in the signaling of narrative structure. In: RANCHHOD, E.; MAMEDE, M. (Org.). *Advances in natural language processing*. Berlin: Springer, 2002b, p.43-51.

\_\_\_\_\_. *Prosodic features in spontaneous narratives*. 2000. 271f. (Thesis) - Simon Fraser University, Vancouver, 2000.

PASSONNEAU, R. J.; LITMAN, D. J. Intention-based segmentation: human reliability and correlation with linguistic cues. In: THE ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 31., 1993. *Proceeding of the ACL 93*, Columbus, 1993. p.148-155.

PICKERING, J. et al. Prosodic markers of salience in humorous narratives. *Discourse processes: a multidisciplinary journal*, Norwood, v.6, n.46, p.517-540, 2009.

PRINCE, G. *Narratology: the form and functioning of narrative*. Amsterdam: Mouton Publishers, 1982.

RUMELHART, D. E. On evaluating story grammars. *Cognitive Science*, Norwood, n.4, p.313-316, 1980.

RYAVE, A. L. On the achievement of a series of stories. In: SCHENKEIN, J. *Studies in the organization of conversational interaction*. New York: Academic Press, 1978. p.113-132.

SABIN, E. J. *Pause and rate phenomena in adult narratives*. 1976. 78f. (Thesis) - Saint Louis University, Saint Louis, 1976.

SABIN, E.; CLEMMER, E. J.; O'CONNELL, D. C.; KOWAL, S. A pausological approach to speech development. In: SIEGMAN, A.; FELDSTEIN, S. *Of Time and Speech*. Hillsdale: Lawrence Erlbaum, 1979. p.35-55.

SACKS, H. Lecture notes: stories in conversation. In: JEFFERSON, G. *Lectures in Conversation*. Oxford; Cambridge: Blackwell, 1972.

SACKS, H.; SCHEGLOFF, E. A.; JEFFERSON, G. The preference for self-correction in the organisation of repair in conversation. *Language*, Baltimore, n.53, p.361-382, 1977.

SCHIFFRIN, D. *Discourse Markers*. Cambridge: Cambridge University Press, 1987.

SCHIFFRIN, D. How a story says what it means and does. *Text: an interdisciplinary journal for the the study of discourse*, Berlin, n.4, p.313-346, 1994.

SELTING, M. On the interplay of syntax and prosody in the constitution of turn-constructural units and turns in conversation. *Pragmatics*, San Diego, v.6, n.6, p.371-388, 1996.

\_\_\_\_\_. Prosody in conversational questions. *Journal of Pragmatics*, Amsterdam, n.17, p.315-345, 1992.

SILVERMAN, K. *Natural prosody for synthetic speech*. Cambridge: Cambridge University Press, 1987.

SWERTS, M. Prosodic features at discourse boundaries of different strenght. *Journal of the Acoustical Society of America*, New York, v.1, n.101, p.514-521, 1997.

SWERTS, M.; GELUYKENS, R. Prosody as a marker of information flow in spoken discourse. *Language and Speech*, Middlesex, n.37, p.21-43, 1994.

UHMANN, S. Contextualizing relevance: on some forms and functions of speech rate changes in everyday conversation. In: AUER, P.; LUZIO, A. *The Contextualization of Language*. Amsterdam: Benjamins, 1992. p.297-336.

VAN DONZEL, M. *Prosodic aspects of information structure in discourse*. Amsterdam: University van Amsterdam Press, 1999.

WOLFSON, N. Speech events and natural speech. *Language in Society*, Cambridge, n.5, p.189-209, 1979.

WOOD, S. Speech tempo. *Working Papers of the Phonetics Laboratory*, Lund, n.9, p.99-147, 1975.

Recebido em 29 de setembro de 2011.

Aprovado em 20 de agosto de 2012.

