PROSODIC FEATURES IN SPONTANEOUS NARRATIVES

by

Miguel Oliveira, Jr.

B.A., Universidade Federal de Pernambuco, 1992

M.A., Universidade Federal de Pernambuco, 1995

THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS OF THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the

Department of Linguistics

SIMON FRASER UNIVERSITY

November 2000

Abstract

Any type of discourse consists of spans of utterances (*segments*) displaying a coherent relationship. Narratives are thought to be composed of clearly recognizable semantically independent segments (or *sections*). There exist several approaches attempting to describe this underlying grammar (or *model*) of narrative. However, no empirical evidence of such a grammar, using spontaneous, non-elicited data has been presented so far.

The present study examines the role prosody plays in the assigning of boundaries in narrative discourse. It also investigates the recurrence of specific prosodic features in semantically individualized narrative sections. If prosody is found to reveal the segmentation of narrative and to characterize narrative sections by means of a recurrent pattern, this constitutes evidence that (oral) narrative presents a well-defined structure.

The data consist of 17 spontaneous / non-elicited narrative texts, told in the course of conversations between the researcher and 8 subjects, under appropriate laboratory condition.

Inspired by the literature, the following prosodic variables were selected for analysis: pause, speech rate, pitch range, pitch reset and boundary tones. It was found that the boundaries occurring between narrative sections (narrative boundaries) are prosodically different from those that occur exclusively between clauses (clause boundaries). This was verified in terms of (i) pause occurrence and duration (pauses occur more frequently and are generally longer at narrative boundaries), (ii) pitch reset

values (the difference in pitch range values between two adjacent clauses is higher at narrative boundaries) and (iii) boundary tones (low boundary tones usually occur at narrative boundaries). Further, it was observed that narrative sections are regularly characterized by the recurrence of specific prosodic patterns, such as (i) variation in speech rate (slower speech rates characterize sections considered highly relevant, faster rates characterize sections of low relevance), and (ii) variation in pitch range (higher pitch ranges characterize sections considered highly relevant, lower pitch ranges characterize sections of low relevance).

These findings lead to the conclusion that spontaneously produced narratives are framed by an underlying structure, and that storytellers appear to be aware of such a structure. This awareness is evidenced through the systematic employment of prosodic devices.

Acknowledgements

This thesis was only possible because of the help and support of a number of people. I would like to thank everybody who directly or indirectly contributed to it. Some people deserve special mention.

Dr. Zita McRobbie, for her inestimable dedication and constructive guidance all the way through the preparation and writing of this thesis.

My colleagues at the Department of Linguistics of Simon Fraser University, for helpful discussion on several technical and theoretical issues. Special thanks to Kyung-Sook Chung, Lorna Marie Fadden, Kyoung-Ja Lee, Herman Li, Tara Jeannine McFarlane, Caroline Morgan, Krista Muller, Grisel Garcia Perez, Dr. Wyn Roberts, Connie So, Dr. Juan Sosa, and Jean Wang.

The staff at the Department of Linguistics of Simon Fraser University, for their help on several aspects of my stay as a graduate student. Gladys Durksen, Carol Jackson, Sheilagh McDonald, and Rita Parmar.

Dr. Rebecca J. Passonneau, Dr. Ester M. Scarpa, and Dr. Marc Swerts, for comments and important suggestions on several aspects of early stages of this research. Dr. Murray Munro and Dr. Nancy Hedberg, for suggestions that have improved the final version of this thesis.

All the Brazilian friends who participated anonymously as informants of this study.

Stella Galindo, Odileiz Santos, Demerval Salviano, Regina Trindade, and GilbertoValadares, for their technical support as Brazilian intonation experts. Reinilda

Freire, Damián Keller, Ana Lúcia Keller, Verônica Oliveira, Graça Salgueiro, Tininha Salgueiro, and Roberto Souza, for acting as trained labellers. Marcos Bagno and Manoel Corrêa, for their help as discourse analysis experts.

Lorna Marie Fadden, for her invaluable help as editor. Tara Jeannine McFarlane, for revising my English.

Simon Fraser University, the Department of Linguistics, and the Donor of the Dr. Tai Whan Kim Memorial Graduate Scholarship in Language and Linguistics, for financial support that helped me to conclude this work.

My parents, for all the encouragement I have received throughout my life. My sisters and brother, for their indispensable support.

Verônica and Lucas, for their love.

Table of Contents

List of Tables

List of Figures

Chapter 1: General Introduction

It is impossible to combine (to produce) a narrative
without reference to an implicit system of units and rules

Roland Barthes,
Introduction to the Structural Analysis of Narratives

Abstract

The purpose of this introductory chapter is to put forward the general goals of the present dissertation, by discussing a few important theoretical concepts that are relevant to the comprehension of where, in the current debate on the relation between discourse structure and prosody, this study stands for, and what contribution it aims at providing. The chapter commences with a brief and general review of the literature encompassing the broad study of narrative analysis, and with the presentation of some basic concepts of prosody. Following this, the objectives of the thesis are established, with the enumeration of the hypotheses to be tested. The chapter also describes the empirical database, the collection procedure and the methods used to prepare the material for the analyses. It concludes with an outline of the thesis as a whole.

1.1 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"[1]. 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 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 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" (Swerts & Geluykens 1994), "information units" (Geluykens & Swerts 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; Reichman 1985; Webber 1988), cue phrases (Hirschberg & Litman 1993; Cohen 1984; Passonneau & Litman 1993), discourse markers (Schiffrin 1987; Hansson 1999; Horne *et al.* 1999), reference (Webber 1988; Grosz & Sidner 1986; Linde 1979) and tense (Hwang & Schubert 1992; Song & Cohen 1991; Webber 1988).

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 1992; Collier, Piyper & Sanderman 1993; Litman & Passonneau 1995; among others), speech rate (Lehiste 1982; Koopmans-van Beinum &

[1] Leech & Short (1981: 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 the present work, the words "discourse" and "text" will be used interchangeably.

Van Donzel 1996; Fon 1999; Selting 1992), and amplitude (Brown, Currie & Kenworthy 1980; Hirschberg, Nakatani & Grosz 1995; Hirschberg & Grosz 1992; Grosz & Hirchberg 1992) 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)[2] 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; Kintsch 1974; Kintsch & Greene 1978; Stein & Glenn 1979). Several approaches to describing this underlying grammar (or model) of narrative discourse have been proposed (Barthes 1975; Black & Wlilensky 1979; Brémond 1973; Genette 1972; Greimas 1970; Labov & Waletzky 1967; Labov 1972; Mandler & Johnson 1977; Prince 1982; Rumelhart 1975; Rumelhart 1977; Rumelhart 1980; Thorndyke 1977; Todorov 1970). Though most of these works are suggestive, there is still very little empirical evidence for any proposed narrative structure.[3]

The present dissertation investigates the role of prosodic features, such as pausal phenomena, speech rate and pitch, in making the structure of spontaneous, non-elicited narrative texts explicit. It will be suggested here that prosodic elements can provide fundamental insight into the way narratives are actually organized, and therefore, serve as

[2] These words will be used interchangeably throughout this study.
[3] See, however, Gee & Kegl (1983), Gee & Grosjean (1984), and Gee (1986) for an approach.

a paradigm for both the building of new theories and the substantiation of existing ones in the broad field of "narratology" (Todorov 1969; Prince 1982; Bal 1985; Chatman 1988).

This opening chapter contains a brief introductory overview of some of the major concepts that constitute the theoretical foundation of the present investigation. It starts with a look into the broad area of research known as "Narrative Analysis" (Section 1.2). A working definition of narrative will be proposed and used as one of the criteria for the selection of the empirical material of the present investigation. This section also addresses two important aspects in the study of narratives that are fundamental for the present research: the structure of narrative texts, as viewed from different perspectives, and the question of narrative expressivity, which has been the central focus of investigation in narrative analysis along the lines proposed by Labov & Waletzky (1967) and Labov (1972).

Section 1.3 contains a brief overview of the second major topic underlying the present research: prosody. It focuses on those aspects of prosody that are closely related to the present work. A definition of prosody is presented, along with a concise presentation of the three acoustic components that are often associated with prosodic studies (duration, pitch and loudness). The notion of "intonation unit" as the basic domain of prosody is also discussed.

The goals of the present dissertation will be outlined in Section 1.4. Besides discussing one of the most problematic drawbacks in the study of prosody as a cue for discourse segmentation, and providing a description of few methodological measures that are often taken to overcome this problem, this section also presents the hypotheses of the present dissertation.

Section 1.5 contains detailed information on the empirical material that is used in the present investigation. It provides a brief discussion of the type of (spontaneous and non-elicited) data that was chosen to form the database of the study. In addition, it presents the methods that were employed for the collection of the corpus. Information about the participants of the project, as well as their individual contribution will be given. Further, the methodological details involving the preparation of the material will be described.

The last section of this chapter, Section 1.6, details the outline of the dissertation as a whole, including a brief summary of each individual chapter.

1.2 Narrative Analysis

The importance of narrative discourse for the understanding of several aspects of our world is incontestable. According to Polkinghorne (1988: 11), narratives are "the primary scheme by means of which human existence is rendered meaningful." Narratives are considered by many scholars to be crucial to the interpretation and reconstitution of cultural, social and personal reality (see, for example, Bruner 1986, Fisher 1987; Rosen 1988). Further, a narrative is an omnipresent component of conversation. As various conversational analysts have pointed out, conversations are often structured around stories (Kirschenblatt-Gimblett 1974; Darnell 1974)[4]. Not surprisingly then, narrative analyses have been the focus of investigation in several areas of research, such as literary sciences, theology, history, psychology, anthropology and linguistics.

[4] For the purpose of this study, the terms "narrative" and "story" will be used interchangeably. In the literature, this is not often the case, though (see Fleischman 1990: 106-107 and Polanyi 1982:511 for a more detailed discussion on this matter).

1.2.1 Definition of narrative

The definitions of narrative are numerous and often controversial (see Oliveira 1995 for a brief discussion). For the purposes of this study, a narrative will be considered to be the speaker's encoding of past experiences into a perceived sequence of discrete, temporally and non-randomly ordered units that accomplishes a *point*.[5]

Narrative texts always make reference to actions that take place in the past relative to the time of the narration, even if they refer to imaginary futuristic actions, such as works of science fiction (Gülich & Quasthoff 1985; Polanyi 1985; Toolan 1988). These actions undergo some sort of transformation or change that may be not only spatial, but, more crucially, temporal (Van Dijk 1972; Brémond 1973; Propp 1958; Todorov 1977).

The fundamental temporal organization system of narrative texts has been widely acknowledged in narrative studies as a basic truism. Starting with Aristotle, and underlying, for example, the Russian Formalists' discussion of fabula and *sjuzhet*, or the French Structuralists' analysis of story and discourse, temporality (or causality) has been regarded as the defining property of narrative. However, it is only in recent discourse-oriented studies that narrative texts have been actually defined by properties of their temporal organization.

According to Labov (1972: 359), a minimal narrative is "a sequence of two clauses which are temporally ordered." These clauses must present a "temporal juncture":

---

[5] The *point* of a story can be defined as "a reason for being told that goes beyond or is independent of any need for the reporting of events" (Johnstone 1990: 18).

the order of presentation in the text (the *sjuzhet* or discourse) must be identical to the order of the events in the depicted world (the fabula or story).[6] In other words, narrative clauses must occur in a fixed presentational sequence in order to retain the original semantic interpretation of a story. The reordering of such clauses would not be impossible or constitute a violation of the conditioning factors of narrative texts, but would result in the telling of a different story. The establishment of such state of affairs is, however, the prerogative of the audience, as suggested above.

In order to clarify the point being made, consider the following example: "imagine you enter a cartoonist's studio and find three frames, on separate pieces of paper, on his desk. They have quite different characters, settings, furniture, etc., and seem to be about quite unrelated topics. ... You think you see a narrative before you — though the cartoonist hotly denies this!" (Toolan 1988: 8). Toolan argues that this example touches on a fundamental feature of narrative analysis: the indispensable role of the story's audience. According to him, "it is idle for anyone else (e.g., the teller) to insist that here is a narrative if the addressee just doesn't see it as one." Although it is the task of a storyteller to provide as much information as needed in a coherent and organized way as to make her/his story meaningful, it is up to the audience to "perform" these meanings (Bruner 1986: 25). Genette (1980: 32) claims, "the real author of the narrative is not only he who tells it, but also, at times even more he who hears it." The narrative text is therefore always the result of a jointly created process.

In that way, the definition proposed above also attempts to capture the dynamic relationship of the narrative discourse with the context of its occurrence and its potential

[6] This definition differs radically from the definitions in literary criticism, which considers temporality a defining property of the fabula, rather than the sjuzhet.

addressees, something that has been often neglected. Narratives create worlds that are pertinent to a socio-cultural community, to its members, and to a specific conversational context. Therefore, narratives are expected to be *locally* and *globally* relevant texts (Johnstone1990: 18).

As mentioned above, narratives often display a transformation or change in the course of the events or actions reported. This transformation or change generally conveys some sort of conflict with expectations. Predictable events are, most of the time, not considered to be an appropriate source for storytelling. As to illustrate this argument, consider the following: "John went to a restaurant. He asked the waitress for coq au vin. He paid the check and left" (Schank & Ableson 1977: 38). Although this series of events could be regarded as a narrative from a formal standpoint, in a casual conversation this would hardly be the case.[7]

What makes this narrative inappropriate is its lack of a manifest point. Stories without a point are condemned to failure. This would lead to another important characteristic of everyday natural narratives: its *reportability* (Hymes 1981; Labov 1972). Stories are told to convey a message in a particular situation; they are intended to make a point. A failure to do so would constitute a violation of a shared conversational expectation and would most certainly give rise to the question that most storytellers are continually warding off: "So what?" (Labov 1972: 366).

It is important, however, to make a clear distinction between the point of telling a story and the point of the actual story. According to Shuman (1986: 81), the former constitutes the story's *storyability* and the later its *tellability*. This distinction allows for

[7] Note however that in other situations, such as a police report or a courtroom testimony, this would be exactly what would be expected (Walker 1982). See Polanyi (1982: 515) for a discussion of the distinction between "stories" and "reports."

an event to be storyable in its uniqueness and yet not tellable in a particular situation and vice versa (see also Johnstone 1990; Robinson 1981).[8] Furthermore, as Polanyi (1979: 207) observed narrative reportability is socially and culturally constrained. More specifically, what might be regarded as reportable in a given socio-cultural context might not hold true in others.

A number of other formal aspects are also taken into consideration for the definition of narrative in the literature. For example, there is a requirement that the participants of the reported events are animate and possess human qualities (Gülich & Quasthoff 1985; Van Dijk 1972). There is also a need for ficcionality in any sort of story. This ensures that the participants of the reported events are unique and not subject to confusion with real people in the real world (Oliveira 1995; Ong 1982; Polanyi 1982). There is also a degree of artificial fabrication and prefabrication, which results in a certain typicality of most narrative texts (Toolan 1988). Finally, there is the display of a linguistic macrostructure which is most of the time regarded as an underlying expectation of narrative trajectory: beginnings, middles, and ends, as stipulated by Aristotle in his *Art of Poetry* (Gülich & Quasthoff 1985; Labov 1972; Labov & Waletzky 1967; Polanyi 1985).

Although many of these aspects are disregarded in different studies carried out in diverse fields of research, one in particular appears to be considered in all studies of narratives: the underlying structure of narrative texts.

---

[8] Johnstone (1990: 19) claims that pointless stories can be perceived as stories sometimes (as, for example, "Shaggy dog stories"). The fact that stories such as those are considered to be stories, however, has much more to do with its *tellability* than with its *storyability*.

1.2.2 Narrative structure

The first scholar who aimed to describe the "system of units and rules" that narratives are supposed to have (Barthes 1977) was the Russian folklorist Vladimir Propp (1969, original work published 1928). Propp's theory of story structure is an inventory of all and only the fundamental events (which he calls 'functions') that he identified in his corpus, based on 115 Russian fairytales. According to him, both the number and sequence of these functions are fixed: there are 31 functions, and they always appear in the same sequence.

After Propp's work, several other approaches aiming at describing the underlying grammar of narrative texts 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; Brémond 1973; Genette 1972; Greimas 1970; Prince 1982; Todorov 1970). 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 1975; Rumelhart 1980; Rumelhart 1977; Stein & Glenn 1979; Thorndyke 1977). 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; Kallmeyer & Schütze 1977; Ryave 1978; Sacks 1971; Sacks 1972).

One of the most influential narrative models in linguistics research is that of Labov & Waletzky (1967) and Labov (1972). 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). However, Labov indicates that most narratives do not in fact contain all of these elements. Instead he suggests that only the "complicating action" is necessary for a minimal narrative, since a minimal definition of narrative involves simply a pair of temporally ordered events. However, each of these elements may contribute to the understanding and appreciation of the audience in a variety of meaningful ways, and therefore each ultimately enhances the success of 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. It is an optional element, though a very important one, since it also functions as a request for the extended turn in the course of a conversation necessary to tell the story and also as an advertisement for the narrative: a way to make exaggerations upon that which follows (Toolan 1988).

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. Syntactically, it comprises past

progressive clauses that do not make use of extensive verbs (i.e., verbs involving action). Labov points out that one of the most interesting aspects of orientation is that sometimes its components are strategically delayed, thus creating an effect of surprise.

The **Complication** consists of a series of narrative clauses in the simple past tense[9] that answers the question: "then, what happened?" It is the backbone of the story and builds up to its climax, which is often resolved with the Result. Commonly, the Complication presents a crisis, a problem, a turning point, or, at least, an event of interest. It is occasionally an extended section in a narrative.

The **Result** contains the resolution to a conflict in the narrative. It usually contains free clauses, which began the complicating action.

The **Coda**, like the Abstract, is an optional element in a narrative — it signals the "sealing off" of a story, by returning the listeners to the present moment. A Coda can take the form of a general observation, explicating the effects of the events on the narrator, or the form of a formulaic expression, such as 'that was that'. It is used as a device to reinstate the conversational mode.

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. According to Labov & Waletzky (1967: 37), it "reveals the attitude of the narrator towards the narrative by emphasizing the relative importance of some narratives as opposed to others," thus being a signal as to how the teller intends the audience to receive the story. The Evaluation 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.

---

[9] The simple past is the unmarked tense in several languages for the displaying of discrete events that took place in the past. Other tenses can also be used in a narrative clause, such as the present tense (displaying the so-called historical present). In this case, however, the use is most often evaluative (Schiffrin 1981; Wolfson 1982).

1.2.3 Evaluation: Narrative expressivity

According to Labov & Waletzky (1967), storytelling has a twofold function. The first is a *referential* function that gives the audience information through the narrator's recapitulation of experience in an ordered set of clauses. The second is an *evaluative* function, which serves to communicate to the audience the meaning of the narrative by establishing some point of personal involvement. This twofold distinction has already been proposed by Malinowski (1923: 149), according to whom the "social and emotive function" (the Labovian evaluative function) dominates the referential function.

Labov (1972) provides an extensive list of devices that may be used evaluatively in a narrative.[10] These devices will often display syntactic complexity where the narrative clauses are, in essence, syntactically simple. In principle, any element can act evaluatively, by drawing attention to itself, by being linguistically marked. The various types of evaluation in Labov's model are distinguished according to whether they occur outside or inside the narrative clause.

Evaluations that appear outside the fixed position clauses of narrative, **External Evaluations**, are of five sub-types: (1) the narrative is interrupted while the teller says explicitly what the point of the story is; (2) a comment made by the teller-as-participant at the time of the events is reported; (3) a comment made by the teller-as-participant to another participant of the events is reported; (4) a comment made by any other participant of the narrative is reported; (5) what the participants of the narrative did — and not what they said — is reported.

---

[10] See also Peterson & McCabe (1983) and Caffi & Janney (1994).

Evaluations that are embedded in the narrative texture, **Internal Evaluations**, are categorized into four sub-types: (1) intensifiers, which emphasize a particular event among a series of narrative events: modifiers (adjectives, adverbs); quantifiers (adverbs); wh-questions (why! where!); repeated lexical items; gestures heightened stress; vowel lengthening; and a wide variation in intonation and pitch range; (2) comparators, which refer to events that did not occur, but that might have occurred. The main type of comparators are expressions of negation, modality and modulation, questions embedded in the action; use of or-clauses; imperatives; future tenses; comparative and superlatives; (3) correlatives, which brings together two events conjoining them into a single independent clause: progressives; appended participles; double appositives; and double attributes; (4) explicatives, which are appended subordinate clauses that qualify, or give reason for, the main events reported: for instance, clauses introduced by *although*, *because*, *while*, *since* and so on.

It should be pointed out, however, that the above evaluative devices are not to be considered absolute. As Polanyi (1985: 14) observes, "any device available for evaluation can be used non-evaluatively as well or can be so over-used that it becomes a textual norm." Another point worthy of note is that the display of expressivity in narratives is inextricably bound with sociocultural contexts: the way people evaluate their stories reflects their social and cultural background (Gee 1985; Gee 1989; Heath 1982; Heath 1983; Johnstone 1990; Michaels 1981; Michaels 1991; Michaels & Collins 1984; Nichols 1989; Scollon & Scollon 1981).[11]

From a broader perspective, the Evaluation Model could be considered a systematization of the meta-structure that goes beyond a story's referential function. This

[11] Johnstone (1996) suggests that individual characteristics should also be taken into consideration.

model then appears to be parallel to the Russian formalists' distinction fabula/sjuzhet and to the French structuralists' story/discourse. Since its introduction, it has been refereed to and applied in several disciplines, such as anthropology (Watson 1972), in literary analysis (Carter & Simpson 1982; Maclean 1988; Pratt 1977), in mass communications (Van Dijk 1984; Van Dijk 1988; Van Dijk 1984), in education (Taylor 1986; Wilkinson 1986); and in developmental psycholinguistics (Kernan 1977; Peterson & McCabe 1983).

The notion of "evaluation" has been refined, extended, and operationalized in order to comply with a wide variety of data (Bamberg & Damrad-Frye 1991; Hicks 1990; Hicks 1991; Hudson & Shapiro 1991; Kernan 1977; Peterson & McCabe 1983; Polanyi 1985; Polanyi 1982; Polanyi 1981; Polanyi 1979; Polanyi 1985; Polanyi 1981; Umiker-Sebeok 1979). Alternatively, studies on expressivity in narrative have introduced concepts different from the Labovian Evaluation Model, which nonetheless, resemble it.

For example, the notion of "performance devices" (historical present, gestures, expressive sounds, repetition, direct speech and asides), as introduced by Wolfson (1979, 1982), is closely related to the Labovian conception of "evaluation." Tannen's (1979, 1982, 1983, 1985, 1989) "involvement strategies" (stylistic devices which display the discourse in a way that transmits the teller's attitudes and feelings towards the events narrated and shapes the listener's engagement in them: repetition, formulaic expressions, ellipsis, direct speech, expressive phonology etc.) and "evidence of structures of expectations" (linguistic elements shaped by culturally-determined beliefs and attitudes: negative expressions, moral judgments, repetitions, contrastive connectives, generalizations, hedges, etc.) are also comparable to the notion of "evaluation," since

most of the devices which are used to convey these concepts are also incorporated into Labov's typology of evaluative categories.

The Evaluative Model has generated some controversy in a wide range of different disciplines. In cognitive psychology, the Labovian model is viewed as surface oriented, since it does not make any reference to cognitive structures or processes (Bamberg & Damrad-Frye 1991). In conversational analysis, Labov's work is criticized because of its disassociation with the social organization of conversation (Gülich & Quasthoff 1985; Gülich & Quasthoff 1986). In anthropology, the Evaluation Model is regarded as being culture specific, since Labov has only correlated narrative functions and structures with age, social class and ethnicity in the United States (Brewer 1985; Zhang & Sang 1986). In literary criticism, Labov's definition of narrative is considered too strict. The criterion of temporal ordering of clauses rules out flashbacks, flashforwards, embedding and subordination, which are common features in literary narratives (Toolan 1988).

1.3 Prosody

Prosody may be defined as "the systematic organization of various linguistic units into an utterance or a coherent group of utterances in the process of speech production" (Fujisaki 1997: 28). It often refers to any of the following auditory aspects of speech: duration, pitch, and loudness.

1.3.1 The intonation unit

Underlying many descriptions of prosody[12] is the idea that there exists a basic structural unit to which prosodic phenomena are anchored. Several approaches that aim to establish the domain of prosody have been proposed, but in fact this makes the task of trying to describe and compare different accounts very complex. In general terms, the considerations that are often made in the classification of the basic prosodic unit are physiological, semantic and formal phonetic/phonological. This section will address some of the major assumptions concerning the basic structural domain of prosody, under these different approaches.

One of the earliest definitions of what a prosodic unit might be dates back to Sweet (1906: 45), who suggests that speakers "are unable to utter more than a certain number of sounds in succession without renewing the stock of air in the lungs." He proposes then that language should be divided into *breath-groups*. This physiological account of prosodic domain has garnered support by many scholars later on, such as Lieberman (1967), who proposes that the production and the perception of intonation is closely associated to the interaction of the respiratory and the laryngeal muscular systems.

The analytical unit in prosodic studies has also been frequently defined in terms of the information content conveyed in a given chunk of structure. The term "sense-group" is proposed by Kingdon (1958: 162) to mean "groups of words that have a semantic and grammatical unit – not necessarily complete." This unit has been taken as

[12] For a review from different perspectives, see Couper-Kuhlen 1986, Ladd & Cutler 1983, and Hirst & di Cristo 1998.

the most immediate correlate of the structural unit of intonation (Armstrong & Ward 1931; Bolinger 1989). Another term that bears basically the same concept as sense-group is proposed by Halliday (1985: 275), the "information unit". According to him, an information unit may be defined as a "process of interaction between what is already known or predictable and what is new or unpredictable." Note that, in general, these concepts correspond roughly to the notion of "clause."[13] Halliday (1970: 3) argues that a clause is generally coterminous with an intonation group "unless there is *good reason* for it to be otherwise" (see also Gee & Grosjean 1983; Selkirk 1984; Marcus & Hindle 1985; Steedman 1990 for a similar argument). However, several studies have already demonstrated that this not often the case (Altenberg 1987; Crystal 1969; Cruttenden 1997).

Since most accounts based on physiological, semantic and/or grammatical considerations present some serious limitations, a phonetic/phonological account has been proposed, as an alternative. According to Crystal (1969: 206), "phonological criteria suffice to indicate unambiguously where a tone-unit boundary should go in connected speech in the vast majority of cases." He proposes the existence of a unit composed of at least one prominent syllable with a major pitch movement (the nucleus) and surrounded by boundaries that can be manifested by many different acoustic cues. In her classic study on the phonology of English intonation, Pierrehumbert (1980) suggests that what delimits the "intonational phrase," are boundary tones, which can be either high or low.[14] This approach has also been the target of criticism: and in fact many analysts reject the

[13] See Quirk *et al.* 1972, for a definition of clause and a more detailed discussion on the issue.
[14] A more detailed description on Pierrehumbert's "autosegmental approach" can be found in Chapter 4, Section 4.1.2.

possibility of straightforwardly identifying intonation groups in speech (Brown, Currie & Kenworthy 1980).

According to Cruttenden (1997: 29), both phonetic/phonological and grammatical/semantic considerations must be taken into account in the classification on intonation-groups. He proposes that the assignment of intonation-group boundaries should be made *a priori* by taking into account phonetic cues. Furthermore, in some difficult cases, semantic and grammatical cues should also be considered. Notably, several attempts to classify intonation unit boundaries automatically, by means of statistical modeling, have been made, with high success rates (Wang & Hirschberg 1992; Altenberg 1987; Bachenko & Fitzpatrick 1990; Ostendorf *et al.* 1990). Ideally, the results from such studies would provide the necessary information in order to build a computer algorithm that would detect automatically the limits of an intonation unit. Such a self-operating algorithm does not exist as yet, which makes the task of assigning intonation units still considerably subjective.

1.3.2 The prosodic features

As stated above, prosody often refers to duration, pitch, and loudness. It examines how these features operate in the intonation unit (or the basic prosodic domain). In the chapters to follow, these elements will be discussed in detail (with the exception of *loudness*, which was left out of consideration in the present study[15]). Below there is a concise description of each of these features.

[15] The examination of *intensity*, the acoustical correlate of loudness, was not carried out for methodological reasons. Refer to Chapter 5, Section 5.3 for a justification.

1.3.2.1 Duration

The study of duration in prosody refers to any tempo phenomena that occur in spoken language. It includes the variation of speech rate (the relative length of a number of successive syllables), the assessment of the length of a single syllable in different environments, and the measurement of pause at various levels.

1.3.2.2 Pitch

In phonation, the vocal folds vibrate at varying rates over time. The frequency at which the vibration of the vocal folds occurs within a second is taken to be the measure of fundamental frequency. Pitch is the perceptual correlate of fundamental frequency.

1.3.2.3 Loudness

The variation of the air-pressure coming from the lungs produces variation in amplitude (or intensity). Intensity can be defined as the amount of energy that is present in speech. Loudness is the perceptual correlate of intensity.

1.4 Research Goals

The primary aim of the present investigation is to determine whether prosodic features can be used as empirical evidence to validate the claim that narratives present a structure organized around semantically individualized discourse units. A specific model of narrative analysis, the Labovian Evaluative Model, is used for this purpose.

It has been repeatedly claimed that one of the major problems in the study of prosodic correlates of discourse structure is the risk of circularity that investigations of this type often incur (Brown, Currie & Kenworthy 1980; Swerts 1997; Swerts & Geluykens 1994). This is due to the fact that the segmentation of a discourse is not generally uncontroversial, and most of the time prosody is used as a criterion for establishing its structure, which makes the reason for this investigation its own end.

A few attempts to overcome this problem are reported in the literature, such as the use of data of a very specific type, displaying easily segmentable coherent units (Terken 1984; Swerts & Geluykens 1994; Venditti & Swerts 1996), or the employment of an explicit model of discourse analysis as the starting-point for independent evidence of discourse structure (Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Passonneau & Litman 1997; Passonneau & Litman 1993; Litman & Passonneau 1993).

The present study adopts a theory-based approach in order to investigate whether narrative structure can be validated on the basis of prosodic phenomena. However, instead of using a general model of discourse structure, such as the one proposed by Grosz & Sidner (1986), it employs a more specific model, designed exclusively for narratives: the Labovian Evaluative Model.

A corpus of spontaneous, non-elicited narratives was used in order to evaluate the hypothesis that prosody helps in the task of making narrative structure clearer. The narratives were analyzed and segmented by discourse analysts, using the Labovian Model. This was done in order to avoid the above-mentioned danger of circularity (independent evidence). A more detailed description of the data collection and preparation will be given in the Section 1.5.

The working hypothesis is tested by investigating the role of some major prosodic features as indicators of narrative section boundaries. Inspired by the literature, the present study will address the following prosodic variables for the purpose of the phonetic analyses: pause, speech rate and pitch range/reset. As noted in the introductory part of this chapter, most of these prosodic markers are considered important structuring devices.

The prosodic variables under investigation are analyzed on two different levels. Whether they evince the narrative segmentation at the local level, by signaling narrative section boundaries, and whether they characterize a specific narrative section at the global level, by means of a recurrent pattern, will both be examined. The hypotheses to be tested, then, include those involving the function of the prosodic variables at different boundary sites (local level), and those involving the role they play within different narrative sections (global level).

1.4.1 Hypotheses concerning narrative boundaries

- Hypothesis 1: Narrative section boundaries are prosodically different from clause boundaries. This can be verified, specifically, through:
    - Pause: longer pauses tend to occur more often at narrative boundaries than elsewhere;
    - Pitch reset: pitch reset is higher at narrative boundaries than elsewhere; and
    - Boundary tones: low tones tend to occur at the end of narrative boundaries, whereas non-low tones are generally found inside narrative sections.

- Hypothesis 2: The greater the number of prosodic cues associated with a boundary, the higher the chances are of its functioning as a narrative section boundary.

These hypotheses are aimed at relating the three prosodic variables to different types of boundary in a narrative: boundaries that coincide with the end of narrative sections and boundaries that occur within those sections. Support for these hypotheses comes from studies of prosodic correlates of discourse structure (see references in Section 1.1).

1.4.2 Hypotheses concerning narrative sections

- Hypothesis 3: Narrative sections may be associated with the recurrence of specific prosodic phenomena:

- Pause: pauses tend to occur more often and tend to be longer at evaluative sections;
- Speech rate: speech rate may characterize narrative sections. Faster speech, for example, might be associated with orientations and abstracts, which are characterized by their relatively low relevance.
- Pitch range: pitch range may play an important role in the characterization of narrative sections. Evaluation and complication sections, for instance, may be characterized by an overall higher pitch, because of their status as key elements in narratives.

- Hypothesis 4: The content of the narrative may have an effect on the use and distribution of these prosodic variables.

These hypotheses deal with the information status of the concepts behind each narrative section. Justification for these hypotheses comes from a variety of works dealing with the relationship between different aspects of prosody and the informational content of discourse. Some examples are the works on pitch phenomena, by Bolinger (1989), Brazil (1997), Kumpf (1987), and Selting (1992); on pause, by Balkenhol, Hollien & Hollien (1979); and on speech rate, by Uhmann (1992). The validation of Hypothesis 3 will corroborate the hypothesis that narratives can be segmented into semantically distinctive constituents.

1.4.3 Justification of the research

Although the structure of the narrative is generally considered to be a defining characteristic of the narrative discourse (regardless of the chosen theoretical model), no empirical and systematic evidence using spontaneous, non-elicited data has been presented as yet. The results of the present study may provide important evidence to substantiate the assumption that oral narratives indeed present a well-defined structure and that prosody is a key evidence for the existence of this structure.

As for the topic of prosody in discourse, there is already a considerable amount of work available dealing with a restricted number of languages, such as English (see Bolinger 1989; Brazil 1997; Brown, Currie & Kenworthy 1980; Hirschberg & Grosz 1992; Johns-Lewis 1986; Lehiste 1975; Pierrehumbert & Hirschberg 1990, just to name a few), Dutch (see Collier, Piyper & Sanderman 1993; Gussenhoven & Rietveld 1988; Swerts 1997; Swerts, Geluykens & Terken 1992; 't Hart, Collier & Cohen 1990; Terken 1984; Van Donzel 1999, among several others) and German (see Günthner 1997; Müller 1992; Selting 1992; Uhmann 1992, for example). Although there exist a considerable number of studies on different aspects of Brazilian Portuguese prosody, no attempt has been made to correlate prosodic features to discourse or discourse segmentation by experimental means so far (Moraes 1998). This research also contributes to the ongoing study of the role of prosodic phenomena in discourse analysis.

Lastly, it is worthy of mention that this type of study has become increasingly important in recent years with the advent of speech technology, and the results may be

relevant for the development of speech synthesis applications, automatic speech recognition applications, and speech browsing algorithms (Shriberg et al. 2000).

1.5 The data

The purpose of this section is to describe the data that is used for the analyses. It commences with a discussion of the type of speech material that was selected for this study (spontaneous, non-elicited), followed by a description of the methods that were used to collect the data. The material selected for the analysis will be presented in detail in Section 1.5.3. Finally, the methodology involving the procedures in the preparation of the material for the acoustical analyses is outlined.

1.5.1 Why "spontaneous" and "non-elicited"?

It has been often pointed out that while the speech style known as "read-aloud" can be very useful for research dealing with controlled data (for comparative purposes, for example), it fails to account for several linguistic phenomena that are crucially important when one is making assumptions about the "real" spoken language, as used in every day communication (Swerts & Collier 1992; Beckman 1997). For that reason, a growing number of researchers are currently making use of spontaneously uttered speech (cf. the volume published by Sagisaka, Campbell & Higuchi 1997). However, 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). In general, speech is designated 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.

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. Swerts & Collier (1992), for example, consider speech to be "spontaneous" when it presents, among other things, self-corrections and utterances that seem to be "generated on the fly." Therefore, any sort of unprepared speech (i.e., not read) would be potentially classified as spontaneous from this perspective. Fujisaki (1997: 38), who adopts a gradual, rather than binary, classification of speech style, proposes that the main characteristic of spontaneous speech is that it must occur "by its own internal force, motivation, etc., rather than by external ones." From this angle, any sort of preparation, or planning, would decrease the degree of spontaneity in speech. According to Fujisaki (1997: 39), the higher degree in spontaneity in speech is only found in free dialogues, which is characterized by the non-specification of format, topic or task. Beckman (1997: 15) argues along a similar line, proposing that if spontaneous

speech is what the researcher desires, then speech in laboratory setting, which she defines as "a communicative situation where the speaker is cooperating obligingly with the experimenter's purpose," must be avoided. Of course, by choosing not to use speech recorded in a laboratory setting as an example of spontaneous speech, one has to deal with numerous methodological drawbacks, such as the quality of the output and the informant's legal rights.[16] This is not a new debate. It goes back to Labov's (1971) sociolinguistic discussion of the *observer's paradox.*

According to Labov (1971), any scholar who aims at investigating the "vernacular" of a language, i.e., the language of a linguistic group as spoken on a daily basis among friends and family, is faced with the methodological dilemma of having access to it without interfering with the informant's habitat. 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.

As discussed above, narratives are often told to illustrate a point. When a speaker decides to tell a story during a conversation, without being prompted to do so, s/he makes

[16] The fact that a recording session is not taking place in a laboratory does not necessarily signify that the output is "spontaneous," because the presence of a tape-recorder continues to be an extraneous element in the setting. Furthermore, such recordings invariably present undesirable noises, which might interfere with the acoustical analysis. Another complicating factor is that recording someone's speech without previous consent is not considered ethical by most institutions.

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, 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.

1.5.2 Data collection

The participants in this project were selected from among the researcher's friends. Their mother tongue was Brazilian Portuguese, and they participated in the project on a voluntary basis. All of them were living in the Vancouver area at the time of the data collection. Most were (graduate) students, and had been living in Canada for no longer

than 3 years. They came from different Brazilian states, resulting in data displaying an array of regional varieties.[17] No one knew the details of the research or the researcher's specific area of study. They were asked to sign a consent form, as required by policy at Simon Fraser University and procedures for review of ethical considerations arising from research involving human subjects (Policy R20.01).[18] Table 1.1 below highlights relevant information about the participants in the project.

Table 1.1
List of participants in the project, with information on their gender, age and the region of origin (in Brazil)

Participant	Gender	Age	State
01	Female	35	Minas Gerais
02	Female	37	Brasília
03	Female	25	Santa Catarina
04	Female	29	Paraná
05	Male	29	Santa Catarina
06	Male	29	Santa Catarina
07	Female	30	Rio de Janeiro
08	Female	33	Pernambuco

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 (see Table 1.2 below for a translated version of the list). 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.[19]

---

[17] Participant 03 was born in Haifa, Israel, but moved to Brazil when she was 1 year old; participant 04 was born in Lima, Peru, but went to live in Brazil when she was around 1 year old; participant 05 was born in São Paulo, but moved to Santa Catarina when he was only 2 years old.

[18] See http://www.sfu.ca/policies/research/r20-01.htm for further details.

[19] Note that while the major topics of the talks were proposed in order to follow the common procedures adopted in "spontaneous interviews," the narratives that were eventually selected were not imposed by the topics. As it will be explained later, none of the subjects knew that the purpose of the interview was to elicit narratives.

Table 1.2
List of possible topics to be developed by the participants during the spontaneous interview

Travels	Parapsychology	The First Girl/Boyfriend
Cinema	Real Facts	The First Kiss
Television	Extraterrestrial Life	The First Time
Death	Abroad Life	Love
Dreams	Foreign Language	Violence
Family	First Days in Canada	Marriage
Pets	Childhood	Great Delusion
Diseases	Teenage	Plans for the Future
Accidents	Weird Relatives	
Shocking Facts	Traffic Accidents	

Recordings were made in a sound-treated room using a professional cassette recorder (Marantz PMD201) and an unidirectional dynamic microphone (Genexxa Intertan 33-984 DCA), positioned about 15-30 cm from the participants' mouths.[20] 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 (see Wolfson 1976 for a discussion). 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 friends of 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 to tell stories, nor was it suggested in any way that narratives were to be told. Nevertheless, most speakers naturally told at least one story.[21]

[20] It could be said then that while the data reflects quite accurately the "real" spontaneous speech, discounting the limitations imposed by the methods employed for the collection of the data, it is not ecologically valid, in the sense that the speech events did not occur in ecologically neutral settings.
[21] Two participants were later disregarded for the present research because their talk did not contain a narrative.

1.5.3 Description of the data

A total of 17 narratives 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 in Section 1.2.1. 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.

(2) The absence of listener's feedback cues. Only uninterrupted stretches of speech were selected for the purpose of the present analysis.[22]

(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.

(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.

The contribution of the participants for the data was not equal. Table 1.3 below presents information about the data, as a function of each participant's contribution.

[22] Although narratives are the result of a jointly created process, as discussed above (p. 7), the listener's feedback responses, if considered in the present research, would affect the measures of pause, tempo and average Fo values.

Table 1.3
Overview of the data, as a function of the participants in the project: number of narratives and the corresponding total duration (in seconds) and total number or words uttered by each participant

Participant	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*08*
Total Narratives	02	02	02	02	04	01	01	03
Total Duration	118.26	115.77	213.59	70.80	178.48	109.44	63.59	243.81
Total Words	293	405	486	215	689	305	237	618

Participant 05 contributed with more narratives than the others in this project. This particular speaker showed a high degree of spontaneity in his talk, which was permeated by narratives of all types. He also contributed more in terms of number of words. However, in terms of time, participant 08 contributed more than the other participants. Despite the fact that the participant who contributed more narratives for the present research is a male speaker, there appears to be no pattern to the distribution of the data in terms of gender. Out of the 17 narratives in the corpus, 12 were told by female speakers. This represents 74% of the material, in terms of time, and 69%, in terms of words.

Table 1.4 below brings information on the narratives comprising the data analyzed.

Table 1.4
Overview of some characteristics of the narratives in the data: total duration of narratives (in seconds), total number of words, major topic associated with the story and number of the participant who told the story

Narrative Number	*Total Duration*	*Total Number of Words*	*Topic*	*Participant Number*
01	48.54	132	Parapsychology	08
02	31.49	116	Accidents	05
03	54.23	216	Accidents	05
04	47.16	166	Traffic Accidents	05
05	109.44	305	Dreams	06
06	113.78	291	Childhood	08
07	63.59	237	Diseases	07
08	45.6	191	Violence	05
09	52.69	141	Marriage	01

Narrative Number	*Total Duration*	*Total Number of Words*	*Topic*	*Participant Number*
10	150.72	486	Accidents	03
11	66.39	239	The First Boyfriend	02
12	49.01	157	Accidents	04
13	81.49	195	Childhood	08
14	62.87	189	Accidents	03
15	21.79	58	Parapsychology	04
16	49.38	166	The First Boyfriend	02
17	65.57	152	Extraterrestrial Life	01

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 (narrative 10), has half the time that was established as the limit in the selection criteria. The shortest has only 21.79 seconds (narrative 15). The average duration of the narratives in the corpus was 65.51 seconds.[23] "Accidents" rendered more narratives than any other topic in the list.[24]

1.5.4 The preparation of the material

In order to avoid the so-called "risk of circularity," a series of methodological procedures was taken before the actual analyses 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

[23] There were, however, a few long narratives occurring in the recordings. Participant 03, for example, told a 12 minutes story, which was disregarded from the analysis.

[24] It must be pointed out here that the topics listed in Table 1.4 are not necessarily the topics of the narratives themselves. The narratives appeared in the discourse when those topics (or topics related to those major topics) were being developed. So, for example, narrative 11 is not the story of the participant's first boyfriend, but a story that appeared when that particular topic was being developed.

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. 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. In addition to indicating the location of the intonation unit boundaries for all the narratives, these experts were also responsible for the identification of the possible nucleus of each prosodic phrase they identified as well as the type of boundary tone (low or non-low) at the end of all the intonation units they found.[25] Judgments were in general consistent.

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.

As a first step, labelers were instructed to identify the points in each narrative where the speaker had completed one communicative task.[26] Once the segmentation was done, using the speaker's communicative intention as the criterion, subjects were asked to label each unit s/he has found according to the Labovian Model. It was assumed here that

[25] A more detailed description of these concepts and the methods employed in the classification of boundary tones will be given in Chapter 4.

[26] This informal notion of communicative intention is the same used by Passonneau & Litman (1993), Litman & Passonneau (1993) and Passonneau & Litman (1997). It derives from the works by Grosz & Sidner (1986) and Polanyi (1988), to whom segments in discourse have a specific coherent goal. The concept was explained to the labelers in common sense terms and by examples.

trained labelers can segment and correctly characterize narrative sections on the basis of the informational content of such sections only.

The ability of subjects to agree with one another in terms of segmentation was measured using a figure called percent agreement (Gale, Church & Yarowsky 1992).[27] Percent agreement averaged 90%, a result comparable to other segmentation studies.[28] Results derived from Cochran's Q tests (Cochran 1950) indicate that this agreement is significant.

The agreement among subjects with regards to narrative sections was assessed with Kendall's W.[29] 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 analyses.

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)[30], all narratives (transcriptions only) were then given to two experts in discourse analysis, who

---

[27] "Percent agreement" can be defined as the ratio of observed agreements with the majority opinion to possible agreements with the majority opinion. Following Passonneau & Litman (1997), the present study considered the agreement among four or more subjects as the "majority opinion."

[28] In Passonneau & Litman (1997), the percent agreement for seven subjects on 20 narratives ranged from 82% to 92%. Grosz & Hirchberg (1992) and Hirschberg & Grosz (1992), who used trained labelers to indicate the structure of spontaneous or read speech using the model developed by x, found a percent agreement ranging from 74.3% to 95.1% among seven coders on three texts under two different conditions (text plus speech and text alone). Hearst (1993), who asked subjects to place boundaries between paragraphs of running text in order to indicate a change in topic, reported an agreement of greater than 80%.

[29] Kendall's W expresses the degree of association between more than two variables.

[30] I thank Dr. Rebecca J. Passonneau for suggesting the experiment described in stage 3.

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 above-mentioned danger 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.

1.6 The outline of the study

In addition to this introductory chapter (outlined in Section 1.1. above), the present work contains four more chapters. The remaining chapters are organized as follows.

Chapter 2 investigates the role of pause as a boundary marker, as well as the possibility of pausal phenomena being responsible for the characterization of different narrative sections. Pauses are analyzed both at local and global levels. The distribution and function of pauses that apparently do not serve as a segmentation cue are also explored.

Chapter 3 deals with the function of speech rate in narrative discourse as a segmentation cue. Speech rate (measured as the number of syllables in a second) is primarily considered on a global level. Whether differences in rate within narratives sections could function as cues for making the narrative structure more transparent is addressed.

Chapter 4 explores the role of pitch phenomena in narratives, both on the global and the local levels. Three specific variables are examined in this chapter: pitch range,

pitch range reset and boundary-marking pitch movements. Pitch range is examined both as a cue for the characterization of individual narrative sections, and as an indicator of narrative section boundary. Pitch reset and boundary tones are only examined as possible narrative section boundary predictors.

Chapter 5 correlates the results presented in the three previous chapters by way of discussing the role of prosodic features in narrative segmentation. It also presents final conclusions and offers some suggestions for future research.

Chapters 2, 3, and 4 are characterized by a common basic structure. They start with a brief overview of the studies on each specific topic they address, focusing on discourse-oriented works. Later they introduce the methods employed for the analysis of each variable, and examine some methodological issues often discussed in the literature. An overview of the data is provided, containing detailed information regarding the distribution of each prosodic feature under examination. This overview is followed by analysis of the data. The chapters round off with a general discussion of the main findings and some preliminary conclusions.

Chapter 2: Pause

There is no such thing as absolute silence,
Something is always happening that makes a sound

R. P. Blackmur,
The Language of Silence

Abstract

This chapter addresses the prosodic feature of pause and its distribution in narrative in relation to the role it plays in signaling narrative structure. First, an overview of the study of pause in speech is given together with a brief discussion of some of the most serious methodological flaws previous the pausological investigations. Second, the methods employed are described, followed by the presentation of the results and the discussion of the findings. On the one hand, it was found that both pause occurrence and pause duration are reliable markers of narrative section boundaries, confirming the hypothesis that the pattern of pausing in storytelling is a significant speech cue capable of revealing the underlying structure of narrative discourse. On the other hand, it was found that pausing strategies do not characterize individual narrative sections. Besides discussing the relation of pausing with the organization of narrative section, a couple of other issues are also addressed: (i) the distribution of pauses that do not function as a boundary indicator in the data, (ii) some of the major acoustic correlates of pause occurrence and duration, and (iii) the variation of pause to speech ratio within the sections of a narrative. In general, the analyses carried out in this chapter demonstrate that pausal phenomena have a very important role in several aspects of storytelling.

2.1. Introduction

Pauses are primarily considered to be "periods of silence in the speech of a person" (O'Connell & Kowal 1983: 221). Obviously, not all "periods of silence" are pauses and not only "periods of silence" characterize a pause. In order for a "period of silence" to be considered a pause, it must occur between vocalizations. In other words, silent gaps that occur in speech as a result of taciturnity or reticence, for example, are not considered pauses. Additionally, other phenomena that do not necessarily correspond to the definition of pause presented above are sometimes considered to be so in the literature, such as filled pauses (vocal hesitations: "uh," "er," etc.), repeats, false starts, syllabic or vocalic prolongations, discourse markers, etc. (O'Connell & Kowal 1983: 221).

Pauses perform multiple functions. Their occurrence is determined by several factors such as anxiety, emphasis, interruption, intersubjectivity, availability, breathing, syntactic complexity, etc. (O'Connell & Kowal 1983: 221). However, what necessitates pause is of little interest in this study. Rather, pauses are mainly regarded as a linguistic cue for narrative segmentation.

It is generally accepted that among all the prosodic features available to the speaker for signaling structure in a text, pause is one of the most efficient (Gee & Grosjean 1984; Rosenfield 1987; Collier, Piyper & Sanderman 1993). It is expected that at some points in the telling of a story, boundaries are more likely to occur than at other

points (see relevant discussion in Chapter 1, Section 1.2.2). Such boundaries are often realized as a period of silence of varying length, in conjunction with other acoustic cues.

In this chapter it is hypothesized that the segmentation of narrative into sections is systematically evidenced in speech by means of pause occurrence and length — speakers systematically signal the end of a narrative section by producing a pause of long duration. Pauses of shorter duration are, on the other hand, regarded as a cue for non-finality — speakers would use them to indicate that the following information has some sort of semantic connection with the previous one.

Further, it is also hypothesized that narrative sections are characterized by different uses of pausing strategies. If it is true that the content of a discursive unit determines the occurrence and duration of pauses within it, as psycholinguistic research proposes (Goldman-Eisler 1968), it would be expected then that different narrative sections would present varying pausing strategies.

The outline of this chapter is as follows: a brief chronological overview of the study of pause is given (Section 2.1.1). A discussion of the role of pause as a segmentation cue will follow (Section 2.1.2). Section 2.2 addresses some major methodological problems that researchers dealing with the measurement of pause in speech expect to encounter (for example the identification of a cut-off point). Subsequently, the methods to be employed in this research are presented. The next section (2.3) introduces the analysis of the data with an examination of the distribution of pauses in the narratives. Based on the results discussed in this section, Section 2.4 examines whether the primary function of pause is to reveal the segmentation of intonation units, as it is proposed in the literature (Rosenfield 1987). Section 2.5 studies

the occurrence of pauses as a narrative section boundary indicator cue. The duration of pause as an indicator of narrative section boundary is investigated in Section 2.6. Section 2.7 analyzes the distribution of pauses that do not coincide with any type of boundary. The use of stylistic pauses within narratives is examined in Section 2.8. Section 2.9 explores the possibility of correlating pause occurrence and duration with other prosodic features. The hypothesis that argues for a correlation between pausal phenomena and narrative sections is tested in Section 2.10. Section 2.11 examines the variation of pause to speech rate within the sections of a narrative and suggests that the "cognitive rhythm," as proposed by Henderson, Goldman-Eisler & Skarbek (1966), Goldman-Eisler (1967), and Butterworth & Goldman-Eisler (1979) can be verified in narrative texts and that this phenomenon reflects the structure of narratives. This chapter concludes with a general discussion of all the findings and some preliminary conclusions (Section 2.12).

2.1.1. The development of pausological research: a brief overview

The purpose of this subsection is to provide a very brief chronological overview of the historical development of pausological research. The information to be presented here aims primarily at showing the direction of the study of pausing. For a more detailed review on the literature, see O'Connell & Kowal (1983).

One of the first attempts to empirically investigate the use of pause in speech was made by Wallin (1901), who defined this prosodic feature as "gaps which, in the main purposeful, separate group of words by silences other than those involved in the mere production of a series of sounds" (1901: 75). This paper brings the first documented physical measurement of pause in the literature. According to Wallin (1901: 81), words

are grouped into "short unities that agree, in the main, with the 'unity of consciousness'"; pauses would be used then to "easily grasp and synthesize this manifold of sensation."

The follow up to Wallin's investigation is Lote's (1911) study on the oral reading of French poetry. Lote challenged the common belief that pausing is primarily used for physiological reasons (breathing), by proposing that it actually occurs as a function of the comprehension of a certain passage of the poem by the reader. Snell (1918: 13), using similar data (reading of poetry), proposed that pauses function to separate "from each other ideas, in order that the mind may grasp these as logical units." It is interesting to note that even in these early approaches to the study of pause in speech, there was a clear preoccupation with relating pausological phenomena to textual and cognitive functions.

Objective methodology for the study of pause was introduced only in 1935, by Parmenter & Treviño. After conducting measurements based on photographic records of an oscillographic display – something relatively sophisticated for that time, the authors concluded that silent pauses accounted for 24 per cent of the total duration of the discourse under investigation (an anecdotal narrative).

The account of pause as a linguistic cue was first proposed by Pike (1945). He considered pause to be comparable in function to contour, pitch, and rhythm. During the 1950s, two important articles appeared as the most significant influence on pausology: Lounsbury (1954) and Maclay & Osgood (1959). Both studies developed a psycholinguistic-cognitive approach for pausological research that was employed by several other analyses afterwards. In the 1960s, Bernstein (1962) inaugurated a sociolinguistic perspective on pausology, by investigating the relation between social class and pause variables.

The course of the development of pausology was finally determined, however, by the research and theorizing of Goldman-Eisler (1958a, 1958b, 1961a, 1961b, 1961c, 1967, 1968 and 1972).

Goldman-Eisler's studies have been concerned with the relationship between thinking and spontaneous speech. More specifically, they were concerned with the question of whether speech pauses (particularly unfilled ones) reflect difficulties in formulating thoughts and expressing them. According to her findings, there is a significant relationship between the length of a pause and its location. The longest pauses were found at boundaries of linguistic units, rather than within those units. Furthermore, subjects paused more often and longer before words that could not be predicted from the context of the discourse than before those that could be predicted. In a review of her own work, Goldman-Eisler (1968: 26) concludes the findings of her studies with the following words: "pauses represent that aspect of the speech act which has little call on skill and which reflects the non-skill part of the speech process. One might regard pausing as an attribute of spontaneity in the creation of new verbal constructions and structures, i.e., of verbal planning."

Currently pausological research represents a substantial portion of linguistic studies taking place within the general domain of psycholinguistics. Numerous studies have been conducted regarding the function of pauses and other temporal phenomena such as speaking rate, phonation-time ratio, articulation rate, length of runs, etc. (e.g. Barik 1977; Boomer & Dittman 1962; Henderson, Goldman-Eisler & Skarbek 1966; Swerts 1994; and particularly the compilation of articles edited by Siegman & Feldestein 1979 and Dechert & Raupach 1980). The general agreement among those studies is that

pauses, besides reflecting cognitive aspects, play an important role in organizing the speech.

The following subsection will focus on the role of pause as a segmentation clue in speech. It will present a general overview of the literature on this particular topic.

2.1.2. Pause as a segmentation clue

It is generally accepted that pausing is reflected in the syntactic hierarchy of a language. Longer pauses occur most of the time between sentences, rather than between clauses (Ford 1982; Grosjean 1980; Hargreaves & Starkweather 1959; Hawkins 1971; Maclay & Osgood 1959; Stenström 1990; Strangert 1991; Verzeano & Finesinger 1949). What many suggest is that this is due to cognitive processes that are needed in the formulation and perception of linguistic units — longer pauses not only provide the speaker time to formulate the next sentence, but also help the audience to syntactically segment the input[1] (see Ford 1982; Goldman-Eisler 1961; Miller, Galanter & Pribram 1960; Siegman 1979).

However, there is no agreement among scholars that the relationship between pause and syntax is a definite one. Brotherton (1979), for instance, found no isomorphism between pause and clause in her study; Beattie (1983) noted that the planning of fluent units typically transcends clause boundaries (see also Gee & Kegl 1983, Sabin *et al.* 1979 and Pakosz & Flashner 1988). This seems to be particularly true in spontaneous speech. Henderson, Goldman-Eisler & Skarbek (1966), for example, found that in spontaneous

[1] Abrams & Bever (1969) and Green (1977) suggest that listeners do most of their cognitive work at the ends of sentences, what would make sentence boundary pauses extremely helpful to them.

speech only 55 per cent of pauses occurred at grammatical junctures whereas 45 per cent occurred within grammatical units. Pauses in oral reading, however, almost always occurred at grammatical junctures (see also O'Shaughnessy 1992; Deese 1984; Strangert 1993). It has been suggested that a much more consistent scenario is found when semantic, pragmatic and prosodic aspects are taken into consideration (Goldman-Eisler 1972; Chafe 1989).

According to Goldman-Eisler (1972), speakers organize their message into highly cohesive sentences with clear hierarchical structure, referred to as "thought units." Chafe (1989) argues similarly, proposing that speech is organized into clusters of information, or "idea units" that can be active at any one time (see Chapter 1, for a more detailed discussion of this topic). Both authors agree that these informational units are demarcated, most of the time, by longer pauses.

Based on the fact that speakers often segment units of information in speech by means of pause, Brown, Currie & Kenworthy (1980: 57) propose a special designation for such pauses: 'topic pauses'. According to these authors, 'topic pauses' (which would be defined, in terms of length, as any pause between 600 and 800 ms) mark bigger conceptual breaks than shorter pauses (see also Selkirk 1980).

In a series of studies on the underlying phonetic characteristic of paragraphs in English, Lehiste (1975, 1979, 1982) observes that in terms of perception, pause length is one of the main cues. According to her findings, the shortest pause associated with a paragraph boundary is 520 ms.

Although pause duration may be associated with an assortment of uses, such as breathing, marking emphatic stress, hesitating, reflecting situational and dispositional

anxiety, planning, etc., any attempt to separate out these operations has always proved to be difficult (see, for example, Grosjean & Collins 1979). As Grosjean (1980) points out, silent pauses are the product of a number of simultaneous operations. For example, a speaker may use a pause to mark a grammatical boundary but may, at the same time, be using this period of silence to breath or plan. For that reason, a dichotomic distinction between hesitation pause and juncture pause has been proposed (see Boomer & Dittman (1962), for example). According to Boomer & Dittman (1962), pauses that occur after a "terminal juncture" (which would correspond, to some extent, to a prosodic phrase) tend to be perceptibly longer than hesitation pauses.

More recently, several works conducted on prosodic characteristics of discourse segmentation have proposed that pause is indeed one of the most important factors in determining the structure of a discourse, both in terms of production and perception (Collier 1993; Collier, Piyper & Sanderman 1993; Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Hirschberg, Nakatani & Grosz 1995; Passonneau & Litman 1993; Pijper & Sanderman 1994; Swerts 1997, just to name a few). Swerts & Geluykens (1994), for instance, have shown that pause structure and topical organization of discourse maintain a strong relationship. Using data from an instruction monologue, the authors illustrate that pauses encode larger-scale information, thus providing internal coherence to discourse (see also Swerts & Geluykens 1993 and Swerts, Geluykens & Terken).

In an attempt to investigate to what extent discourse boundaries — obtained by means of comparing variances between naïve labelers — correlate with specific prosodic variables, Pijper & Sanderman (1994) and Swerts (1997) found a significant correlation

between boundary strength values[2] and pause duration. According to Collier, Piyper & Sanderman (1993), pause alone shows sufficiently clear relations to the linguistic structure of the utterance.

Grosz & Hirchberg (1992) and Hirschberg & Grosz (1992) found that in a corpus of recordings of Associated Press news texts, phrases beginning discourse segments are correlated with duration of preceding pauses, while phrases ending discourse segments are correlated with subsequent pauses (see also Hirschberg, Nakatani & Grosz 1995 and Nakatani & Hirschberg 1995).

It is noteworthy that although the data in some of these studies are narrative texts (Gee & Grosjean 1984; Gee & Kegl 1983; Grosz & Hirchberg 1992; Kowal & O'Connell 1987; Kowal, O'Connell & Sabin 1975; Kowal, Wiese & O'Connell 1983; Levin, Schaffer & Snow 1982; Litman & Passonneau 1993; Pakosz & Flashner 1988; Passonneau & Litman 1993; Rosenfield 1987), there has not been any attempt to relate the use of pause (or any other prosodic phenomena) to narrative structure, using an independent model of *narrative discourse*[3]. Some analyses (Chafe 1980; Gee & Grosjean 1984; Gee & Kegl 1983; Lehnert 1981; Rosenfield 1987; Scollon & Scollon 1981) propose that pausing indeed reflects the structure of narrative texts, since it marks the boundaries of their units. However, they do not provide a clear explanation of exactly in

---

[2] For a discussion of the notion of *perceptual boundary strength* (PBS), see Pijper & Sanderman (1994).

[3] It could be argued that the studies by Grosz & Hirchberg (1992), Hirschberg & Grosz (1992), Litman & Passonneau (1993), Litman & Passonneau (1995), Passonneau & Litman (1993) and Passonneau & Litman (1997) are a few exception on this particular issue. The first two authors used an explicit model of discourse structure proposed by Grosz & Sidner (1986), while the last two used speaker intention as a criterion for discourse segmentation. However, although these studies made use of a theory-based approach, in an attempt to motivate discourse segmentation based on independent models of discourse analysis, and although they all used narratives as the main data for their analyses, no attempt to correlate prosodic features to the structure of *narrative discourse* specifically was made.

what way pausing contributes to the organization of spontaneous narrative texts from a semantic/pragmatic perspective (see Oliveira 1999 for a treatment of this topic).

The present study aims at taking the hypothesis that pauses play an important role in the organization of narrative texts one step further in order to show how precisely pausing data can reflect the segmentation of spontaneous narratives into different semantic/pragmatic units.

2.2. Methodological issues in the study of pause in speech

It must be pointed out at the beginning of the discussion, that this study is restricted to the analysis of silent pauses, defined by Hieke, Kowal & O'Connell (1983: 203) as "a period of vocal inactivity of a certain duration embedded in the stream of speech." It does not include the occurrence of filled pauses or any other phenomena that include vocalization since they have a very limited incidence in speech (Hieke, Kowal & O'Connell 1983; Brotherton 1979) and are not primarily used as a means of signaling boundaries in discourse (Brotherton 1979; Stenström 1990)[4].

There are two methodological problems related to the location and measurement of pauses in discourse. First, as it has been pointed out by O'Connell & Kowal (1972), the reliance on subjective procedures such as judges' estimation or stopwatch timing of silent

[4] Filled pauses may serve other cognitive and communicative tasks. Fox Tree (1993), for example, suggests that they help listeners to recognize an upcoming target word faster, as most linguistic material followed by a filled pause have a low transitional probability and thus a high information value (Goldman-Eisler 1968). Filled pauses may be used to 'hold the floor' in conversation (Maclay & Osgood 1959; Stenström 1990) or to indicate uncertainty at any point in a talk (Kowal, O'Connell & Sabin 1975; Smith & Clark 1993; Brotherton 1979). Some recent works have suggested that filled pauses may be also used in discourse as a boundary indicator (Chafe 1980; Schegloff 1979; Swerts 1994; Swerts & Geluykens 1994; Swerts, Geluykens & Terken ; Swerts, Wichmann & Beun 1996). However, such use is remarkably peripheral and

pauses is inadequate, due to the resolution capability of the perceptual systems involved and the concomitant errors of reaction time. Second, there is no agreement among researchers concerning the cut-off point for defining a silent pause.

Siegman & Pope (1966) consider a pause to be a silent interval longer than 3 sec in duration. Hargreaves & Starkweather (1959), based on the assumption that human listeners can only perceive pauses of one second and up, set this limit for their experiment. The same criterion was used by Lay & Paivio (1969), Levin & Silverman (1965), and Taylor (1969). Barik (1977) used a cut-off point of 610 ms. Verzeano & Finesinger (1949), without any explanation, set a lower limit for pause at 500 ms. Siegman (1979) opted for a 300 ms limit. Sabin (1976), Sabin *et al.* (1979) and Kowal, O'Connell & Sabin (1975) used a 270 ms cut-off point. Kowal & O'Connell (1987) employed a 230 ms limit for their research, while Boomer (1965) specified a 200 ms cut-off point. Butcher (1981), Henderson, Goldman-Eisler & Skarbeck (1965), Henderson, Goldman-Eisler & Skarbek (1966), and Hieke, Kowal & O'Connell (1983) defined pause as a period of silence equal or greater than 100 ms. A lower limit of 80 ms was used by Levin, Silverman & Ford (1965). Wilkes & Kennedy (1969) used no cut-off point at all. Maclay & Osgood (1959: 24) adopted a perceptual criterion: pauses "were marked when there was judged to be an abnormal hesitation in speech that could not be referred to the three previous categories [repeats, false starts, and filled pauses]." The same principle was employed by Tannenbaum, Williams & Hillier (1965).

The most common cut-off point, however, is up to 250 ms after the precedent of Goldman-Eisler (1961, 1972, 1958, and 1968). According to Goldman-Eisler (1968), the

---

does not constitute per se a sufficient signal of discourse boundary (Stenström 1990; Stenström 1986; Swerts, Wichmann & Beun 1996).

consequences of adopting a minimum cut-off point well above 250 ms are numerous, since 71.5% of all pauses occur in the duration interval between 250 ms and 1 sec (see also Dalton & Hardcastle 1977 for a similar claim). She also points out that the consequences of adopting no minimum cut-off point are serious as well, since short periods of silence — which cannot be considered as psychologically functional pauses — are required for articulation.

Hieke, Kowal & O'Connell (1983), on the other hand, argue for the arbitrariness of the 250 ms cut-off point, by observing that it is not very common to find a pause in the ranges between 130 ms and 250 ms due to articulatory causes. According to them, the articulatory rationale for cut-off points should be 130 ms and not 250 ms, as proposed by Goldman-Eisler (1968) (see also Martin 1970). However, they agree that, for the sake of comparability with existing research data, a traditional (higher) cut-off point should be taken into consideration.

It is important to note that most of the studies mentioned had English as their primary focus of investigation. Some authors propose that although many languages do present a very similar pattern with respect to the use of silent pauses (Black *et al.* 1966; de Johnson, O'Connell & Sabin 1979; Grosjean & Deschamps 1975; Grosjean & Deschamps 1972; Grosjean & Deschamps 1973), cross-linguistic differences may emerge.

Cross-linguistic differences in the study of pause are addressed in a comprehensive study on temporal patterns in six languages (Spanish, Portuguese, French, Italian, English, and German), Glukhov (1975) observed that the distribution of pause across languages varies. For example, according to his findings, the distribution of pause

in Portuguese and Italian differs greatly, both with respect to number of silent pauses per hour and patterning across pause-duration intervals. In this corpus, the Italian subjects used 188 fewer pauses than any other group, while the Portuguese speakers used 225 more pauses.

Regarding the variability in duration of pauses, Glukhov's data displays considerable difference in frequency of occurrence in the interval of 50-150 ms. For example, in his data, the average number of silent pauses per hour in Portuguese was almost twice as great as in Italian (430 vs. 217). As Glukhov points out, this variation is to be expected across languages as a function of the shifts in articulation particular to the sound system of each individual language. On the other hand, the vast majority of silent pauses in all six languages in the study, those in the 150-2000 ms pause duration intervals, did not display a considerable variability. This would not only suggest the existence of a pause duration universal, but would also validate the adoption of a higher cut-off point (Kowal, Wiese & O'Connell 1983).

It should be clear that the study of pausological phenomena cannot be abstracted from their speech context. Information on the type of discourse, social setting and role, age, gender and other variables should always need to be taken into consideration. In a study on temporal patterning in children's narratives, Kowal, O'Connell & Sabin (1975) have found that gender might have an effect on duration of silent pauses: boys generally produce longer pauses than girls (see also Starkweather 1983). Age-educational level may also affect the frequency of pause, as reported by Sabin *et al.* (1979). Bernstein (1962) reported a significant socio-economically based difference between the proportion

of pauses amongst middle-class subjects and lower working-class subjects — the former tended to use more pauses than the latter (see also Brotherton 1979).

Speech style has also a strong influence on pause duration. As several researchers have observed, pause tends to be shorter and less frequent in reading aloud than in spontaneous speech (Barik 1977; Butterworth 1980; Goldman-Eisler 1968; Grosjean & Deschamps 1973). Goldman-Eisler (1968) suggests that the use of pause increases with semantic complexity (as, for example, in a description of a cartoon, that would require time to decode the image and encode it linguistically) and decreases with learning or rehearsal of spoken material.

Interpersonal situations may call for varying frequencies of pause. In an experiment with 10-year-old children, Levin & Silverman (1965) found that they pause more often when telling stories before an audience of adults then when they are alone speaking into a microphone.

Finally, different types of discourse can also affect pause duration and occurrence. According to Barik (1977), lectures and film discussions present an overall shorter pause than public speeches and storytelling. Levin, Schaffer & Snow (1982) also suggest a greater incidence of long pauses in storytelling as compared to reading aloud (see also Sabin *et al.* 1979).

The use of a homogeneous data set is crucial in order to avoid incurring the risk of making generalizations out of different types of material, subjects and speech contexts (Sabin 1976: 5). By using adult participants, belonging to the same social class, the present study offsets this problem. The use of a very specific type of discourse (narratives), collected in the same experimental environment, also contributes to the

homogeneity of the material and makes it possible to compare results with several similar investigations. As Kowal, Wiese & O'Connell (1983: 378) point out, "there are more published experimental investigations of story telling than of any other adequately operationalized speech style."

As for the adoption of a cut-off point — in the absence of strong evidence that pauses of less than 250 ms do not reflect primarily phonational gaps and in order to make sure that the results obtained from this study can be comparable to several other studies that opted for the use of the traditional cut-off point proposed by Goldman-Eisler (see above), only silent pauses of 250 ms or greater between words were taken into account in this analysis. Pauses were measured directly off digitized speech waveforms so as to avoid subjective judgments of location and duration of unfilled pauses (Sabin *et al.* 1979: 36; O'Connell & Kowal 1983: 273).

The following sections will present the results of the analysis, accompanied by a discussion of each issue under investigation.

2.3. The distribution of pauses in the data

Table 2.1 below displays the distribution of pause and speech parts in all seventeen narratives in the corpus, specifying total duration of the narrative as a whole, total duration of speech and pause parts, and pause ratio, broken down for narratives and mean values over narratives.

Table 2.1
Distribution of pause and speech parts: total duration of each narrative, total duration of speech and pause parts, and pause ratio, broken down for narratives and mean values over narratives. The duration values are expressed in seconds.

Narrative	*Duration*	*Speech*	*Pause*	*Pause Ratio*
01	48.54	38.2	10.34	0.21
02	31.49	22.02	9.47	0.30
03	54.23	38.43	15.8	0.29
04	47.16	33.2	13.96	0.30
05	109.44	85.51	23.93	0.22
06	113.78	81.04	32.74	0.29
07	63.59	50.47	13.07	0.21
08	45.6	34	11.6	0.25
09	52.69	39.17	13.52	0.26
10	150.72	122.69	28.03	0.19
11	66.39	57.44	8.95	0.14
12	49.01	38.7	10.31	0.21
13	81.49	56.11	25.38	0.31
14	62.87	49.47	13.4	0.21
15	21.79	16.21	5.58	0.26
16	49.38	42.8	6.58	0.13
17	65.57	50.47	15.1	0.23
Mean	65.51	50.35	15.16	0.24

The results in Table 2.1 confirm the experiments by Parmenter & Treviño (1935) and Brotherton (1979), who found that silent pauses accounted for about 25% of the total duration of speech. In a study based on the works by Duez (1982) and Barik (1979), Johns-Lewis (1986) observed that the speech style might affect the pause to speech ratio. According to her findings, pause ratio is likely to be higher (around 56%) in speech involving "reflective interpretation" than it is in public oratory and interview (20%). It is not very clear what she means exactly by speech requiring "reflective interpretation"[5], but if we consider that the material used in the present study is comparable to an interview (as Wolfson 1976 describes), then the results displayed above concur with the findings of Johns-Lewis (1986).[6]

[5] It could be the case that the author is referring to what Goldman-Eisler 1968 refers to as "the complex cognitive process of interpreting" (see discussion of the topic in section 2.10 below).
[6] See, however, Sabin (1976), who found an average of 35% of silent pauses in his corpus of narrative texts.

The variation in terms of pause to speech ratio across narratives is not high.[7] Most narratives have a ratio between 20% and 30%, with a couple of exceptions. Narratives 11 and 16 have a quite low pause to speech ratio (16% and 13% respectively). It should be pointed out that both narratives were told by the same participant. This may indicate an idiosyncratic characteristic of this person.

Since the above results revealed highly uniform data in terms of the distribution of pause, and since there is already enough evidence that the type of spoken material determines the frequency of pause usage, it follows that the data indeed belongs to a very specific type of discourse. This implies then, that the narratives are safely comparable, on the methodological grounds previously discussed.

The following section displays the results in more detail. Pause distribution is compared as a function of the location where it takes place. The aim of this section is to verify whether pause is mainly used to indicate the end (or the beginning) of an intonation unit, as it is often proposed.

2.4. Pause occurrence as a demarcator of intonation units

It is commonly accepted in any intonation analysis that pause functions primarily to encode intonation unit[8] boundaries (Brown, Currie & Kenworthy 1980; Chafe 1989; Pakosz & Flashner 1988; Selkirk 1980). The role of pauses in discourse has more far reaching effects. The present section will argue that, if considered on a more global perspective, pauses may assume a more particular function in the organization of

[7] A Shapiro-Wilk W Test for normality indicates that the distribution of values is normal ($W < 0.05$).
[8] For a discussion of the notion of "intonation unit," refer to Chapter 1, Section 1.3.4.

narrative texts. First, in order to verify whether intonation units (IU) are consistently segmented by means of pause in the present data, an examination of the proportion of pause occurrence per intonation unit in all narratives is necessary. The result from this analysis is displayed in Table 2.2 below:

Table 2.2
Total number of intonation units and the corresponding numbers of pauses that occur at intonation unit boundaries, broken down by narrative, and percentage values over narratives

Narrative	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*08*	*09*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*Total*
IUs	23	23	42	30	53	56	36	32	34	86	46	26	32	37	14	31	28	629
Pauses	10	11	20	16	19	30	12	15	14	39	15	11	19	24	9	7	13	284
%	43	48	48	53	36	54	33	47	41	45	33	42	59	65	64	23	46	45

There are a total of 629 IUs in all narratives, but only 284 (45%) of them are followed by a pause, a relatively small rate of recurrence, when compared to other studies (see, for example, Geluykens & Swerts 1993; Rosenfield 1987).

Variations in pause occurrence within narratives are also found in the results displayed above (compare, for example, narrative 16 and narrative 14: 23% and 65% respectively). It should be pointed out again that narrative 16 was told by a participant who used a very low speech to pause ratio, and this might be the cause of the small number of pauses at intonation unit boundaries.

If intonation unit boundaries are not systematically marked by means of pausing, as the numbers in Table 2.2 suggests, how should pause occurrence be accounted for? As it will be proposed in Section 2.5 below, narrative section boundary instead might be a better clue for the occurrence of pause in storytelling. The next section will present the distribution of pause in the data in more depth, displaying the values in all the three possible locations where a pause can occur: at narrative and non-narrative boundaries,

and at places that do not correspond to any type of boundary. If the segmentation of the speech into macro units, such as discursive constituents (narrative sections, in this case) is what determines the occurrence of pauses, rather than the segmentation into micro units (IUs), a higher percentage of pause occurrence would be expected at narrative section boundaries than at any other location in a narrative.

2.5. Pause occurrence as a narrative boundary predictor[9]

The results shown in the previous section suggest that pause does not necessarily occur as a function of the existence of an intonation unit. Rather, pausing appears to be a discourse tool that encodes units of a higher level. The present section investigates whether narrative section boundaries systematically incur the occurrence of pause.

Table 2.3 below displays the number of absolute occurrences of pauses at narrative boundaries, at boundaries that do not correspond to narrative boundaries and of pauses found elsewhere in the narratives.[10]

[9] The term "narrative boundary" is used here interchangeably with "narrative section boundary".
[10] The results of pause occurrence at non-boundary sites will be discussed in this chapter.

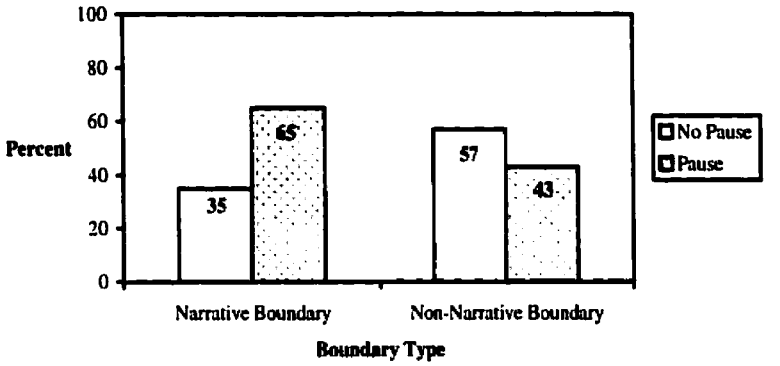
Table 2.3
Distribution of pauses: total number of occurrences (percentages in parenthesis) at narrative boundaries, non-narrative boundaries and elsewhere, broken down for narratives and total values over narratives

Narratives	*Narrative Boundary*	*Non-narrative Boundary*	*Elsewhere*	*Total*
01	4 (36)	6 (55)	1 (9)	11
02	3 (23)	8 (62)	2 (15)	13
03	3 (13)	17 (74)	3 (13)	23
04	5 (29)	11 (65)	1 (6)	17
05	0 (0)	19 (66)	10 (34)	29
06	5 (15)	25 (74)	4 (11)	34
07	2 (11)	10 (53)	7 (36)	19
08	4 (27)	11 (73)	0 (0)	15
09	3 (20)	11 (73)	1 (7)	15
10	8 (18)	31 (67)	7 (15)	46
11	3 (17)	12 (67)	3 (16)	18
12	3 (22)	8 (57)	3 (21)	14
13	6 (27)	13 (59)	3 (14)	22
14	8 (33)	16 (67)	0 (0)	24
15	2 (20)	7 (70)	1 (10)	10
16	4 (45)	3 (33)	2 (22)	9
17	5 (26)	8 (42)	6 (32)	19
Total	68 (20)	216 (64)	54 (16)	338

The values in the above table should be considered with some caution. Apparently, they suggest that pauses occur more often at non-narrative boundaries than at any other location (216 versus 122 pauses). This is true, if one takes into consideration the absolute numbers only. It is also true, however, that the number of boundaries that do not correspond to narrative boundaries in the data is much greater (510 versus 117). In other words, only 19% of all intonation boundaries in the narratives correspond to narrative boundaries. Therefore, the absolute numbers in Table 2.3 do not say much about the distribution of pauses in the data as a function of the type of boundary where they occur.

In order to determine whether the type of boundary determines the occurrence of pause in narrative discourse, the distribution of pauses at narrative and non-narrative boundaries must be considered as a function of the actual number of each type of boundary. The proportion of occurrence and non-occurrence of pauses in narrative and non-narrative boundaries can be found in the pareto chart below (Figure 2.1):

Figure 2.1
Distribution of pauses at narrative and non-narrative boundaries (pareto chart)



It is clear now that a pattern in terms of occurrence of pause in relation to type of boundary does exist in the present data. In 65% of the time that a boundary corresponds to the end/beginning of a narrative section, a pause will occur. If the boundary is not a narrative boundary, this number drops to 43%. This difference is statistically significant (χ^2=16.722, df=609, p<0.0001). Therefore, it's not always the case that an intonation unit boundary will determine the occurrence of a pause, as the literature proposes (see discussion above). The occurrence of pause in discourse seems to be correlated with a higher-level segmentation.

Table 2.4 below gives the total number of narrative boundaries and occurrence of pauses in all seventeen narratives, giving a more detailed look at the distribution of pauses as a function of the type of boundary in the data:

Table 2.4
Total number of narrative boundaries and occurrence of pauses, broken down by narrative, and percentage values over narratives

Narrative	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*08*	*09*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*Total*
Boundaries	4	4	4	7	6	8	4	4	6	10	6	6	8	8	3	8	7	103
Pauses	4	3	3	5	0	5	2	4	3	8	3	3	6	8	2	3	5	67
%	100	75	75	71	0	63	50	100	50	80	50	50	75	100	67	37	71	65

Three narratives in the data (01, 08 and 14) are characterized by the presence of a pause at *all narrative boundaries*, as Table 2.4 above reveals. However, exceptions are also present. Narrative 05, for example, contains no pauses at narrative boundaries. This is interesting, because it completely diverges from the average. Notwithstanding, the distribution of pauses in the non-narrative boundary condition seems to follow a pattern: 59% of the boundaries are not followed by a pause. Narrative 07 also has an odd distribution: 50% of the boundaries do not contain a pause. The distribution in the non-narrative boundary condition follows the pattern, however: 68% of the boundaries are not followed by a pause. Narratives 09, 11 and 12 have a distribution similar to narrative 07. What seems to be pertinent is that with the exception of narrative 07, all the others present a minimum of 50% of pause occurrence in the narrative boundary condition.

In the previous sections, it is demonstrated that pause occurrence is a reliable indicator of narrative boundary. The next step is to investigate whether pause duration also plays a role in the identifying of such boundary.

2.6. Pause duration as a narrative boundary signal

As previously discussed, pause duration is often considered a good indicator of the boundaries between larger discourse units: the longer a pause is, the greater the chances that the place where it occurs coincides with a major discursive break. There are numerous studies confirming this postulation (see, for example, Gee & Grosjean 1984; Gee & Kegl 1983; Grosz & Hirchberg 1992; Kowal & O'Connell 1987; Kowal, O'Connell & Sabin 1975; Kowal, Wiese & O'Connell 1983; Levin, Schaffer & Snow

1982; Litman & Passonneau 1993; Pakosz & Flashner 1988; Passonneau & Litman 1993; Rosenfield 1987, among others). However, as already stated, none of these authors attempt to relate *narrative structure* to pausal phenomena using an independent framework of narrative analysis.

Table 2.4 brings the mean values of duration for each narrative, as a function of the location where they occur.

Table 2.5
Mean pause duration (standard deviation in parenthesis) at narrative boundaries, non-narrative boundaries and elsewhere, broken down for narratives and total values over narratives. The duration values are expressed in seconds

Narratives	*Narrative Boundary*	*Non-narrative Boundary*	*Elsewhere*	*Average*
01	1.32 (0.89)	0.77 (0.33)	0.43 (N/A)	0.94 (0.63)
02	0.49 (0.18)	0.72 (0.27)	0.68 (0.18)	0.66 (0.25)
03	0.79 (0.24)	0.79 (0.30)	0.65 (0.27)	0.77 (0.28)
04	1.12 (0.44)	0.72 (0.38)	0.40 (N/A)	0.82 (0.43)
05	N/A	0.76 (0.43)	0.88 (0.24)	0.80 (0.37)
06	1.57 (0.49)	0.95 (0.68)	0.54 (0.29)	0.99 (0.67)
07	0.98 (0.99)	0.67 (0.24)	0.62 (0.33)	0.68 (0.27)
08	0.81 (0.44)	0.74 (0.30)	N/A	0.76 (0.33)
09	0.97 (0.55)	0.85 (0.44)	0.62 (N/A)	0.86 (0.44)
10	0.63 (0.11)	0.63 (0.32)	0.43 (0.15)	0.60 (0.28)
11	0.35 (0.11)	0.57 (0.28)	0.36 (0.05)	0.50 (0.25)
12	0.86 (0.21)	0.62 (0.26)	0.92 (0.47)	0.74 (0.31)
13	1.32 (0.57)	1.07 (0.78)	1.20 (0.68)	1.15 (0.70)
14	0.58 (0.33)	0.55 (0.33)	N/A	0.56 (0.32)
15	0.87 (0.86)	0.47 (0.20)	0.59 (N/A)	0.56 (0.37)
16	1.00 (0.37)	0.55 (0.03)	0.51 (0.03)	0.74 (0.34)
17	0.92 (0.48)	0.77 (0.32)	0.73 (0.25)	0.79 (0.34)
Average	0.92 (0.51)	0.74 (0.44)	0.68 (0.33)	0.76 (0.44)

The mean duration of pauses occurring at narrative boundaries (0.92 s) differs significantly from those occurring at non-narrative boundaries (0.74 s). T-test result ($t=-2.821$, $df=282$, $p<0.01$) confirms this finding by showing a significant correlation between narrative boundary and longer pause duration.[11]

Again, a considerable amount of variation within narratives can be seen. Narratives 02, 03, 05, 10 and 11 do not follow the general pattern. The mean duration of

[11] The duration of non-boundary pauses will be discussed in Section 2.7.

pauses at narrative and non-narrative boundaries are the same for narratives 03 and 10, while narratives 02 and 11 have a mean pause duration longer at non-narrative boundaries than at narrative boundaries. As discussed above, narrative 05 does not present any pauses at all at narrative boundaries. This is of course an exception.

Interestingly, narrative 05 is also inconsistent with the others in that it displays a high mean duration of pauses at *non-boundary* sites. Only narrative 12 shares this property. It should be pointed out, however, that while narrative 05 presents a large number of pauses at locations that do not correspond to any type of boundary (a total of 10 out of 29), narrative 12 has only a few (a total of 03 out of 14). Nonetheless, two of these pauses are also the longest in the whole narrative.

Analyses of variance were carried out with "pause duration" as a dependent variable and "narrative" and "participant" as fixed factors. Results show significant effects for both "narrative" ($F(16,283)=3.04$, $p<0.0001$) and "participant" ($F(7,283)=6.2866$, $p<0.001$). Post-hoc tests (Tukey-Kramer multiple comparisons) for "narrative" show that narrative 13 differs significantly from narratives 10, 14, 15 and 11 and that narrative 06 also differ from the same ones except for 15. Note that both narratives (13 and 06) were produced by the same participant (08). As for the variable "participant," Tukey-Kramer tests reveal that participant 08 differs significantly from participants 5, 7, 4, 3 and 2, while participant 1 differs significantly from participant 3 only. Table 2.6 below brings the actual mean duration of pause for each participant, along with standard deviations.

Table 2.6
Mean pause duration and the corresponding standard deviation for each participant. Values are expressed in seconds

Narrative	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*08*
Mean duration	0.85	0.61	0.60	0.63	0.77	0.76	0.80	1.06
Sd	0.41	0.32	0.30	0.32	0.33	0.43	0.25	0.68

Participant 8 (narratives 01, 06 and 13) not only has the longest average pause duration, but also the largest standard deviation. Some of the pauses that are employed in the narratives produced by participant 8 were, in relation to the others, exceptionally long, contributing to the large standard deviation displayed in Table 2.6. Long pauses constitute, most of the time, what Glukhov (1975) calls "semantic" pauses. This specific property of pause will be discussed later on in this chapter.

What is important to highlight at this point is that both pause occurrence and pause duration prove to be significant predictors of narrative boundary. Narrative boundaries are characterized by the presence of a pause of longer duration, while non-narrative boundaries do not require the presence of a pause. When they occur, they tend to be shorter than those at narrative boundaries.

In the following section, the distribution of pauses that do not coincide with any type of boundary will be analyzed and discussed.

2.7. Non-boundary pauses

It was shown that pausing (occurrence and duration) in narratives is primarily used to reveal the structure of this particular type of discourse, by segmenting its various sections. Pauses are also used to encode smaller components in narrative texts (as they are in any other kind of discourse), such as intonation units. A small number of the

pauses found in this data however, do not coincide with any of these above-discussed functions. This section will concentrate on such pauses.

Table 2.3 above brings to light an interesting aspect of the distribution of pauses in the data. From all 338 pauses, only 54 (or 16%) are in locations other than IU boundaries. This is a very small number, if one takes into consideration that the material is spontaneous in nature, which implicates, according to the literature, a less restricted distribution of pauses (Barik 1977; Strangert 1993).[12]

Although it is often claimed that the distribution of pauses in spontaneous speech is less restricted than in read-aloud material, they are not distributed haphazardly (Goldman-Eisler 1972; Hansson 1998, 1998; Rosenfield 1987). Most of the pauses that do not coincide with boundary pauses are associated with:

(1) Disruptive utterances:[13]

quando acaba o (0.95) a/ **quando ela dá a volta** (03:05)[14]
when it ends (0.95) ***it/ when it turns back***

eu sonhava com uma/ (1.26) **eu sonhava que eu mexia assim num caldeirão de fogo** (03:07)
I dreamt about a/ (1.26) ***I dreamt that I was stirring something in a big hot pot***

éh duas garotas viram um um ra/ um (0.85) **uma** (0.71) **éh um** (0.79) **um objeto não-identificado** (17:13)
ahn two girls saw a a ra/ a (0.85) ***a*** (0.71) ***ahn a*** (0.79) ***an unknown object***

(2) Discourse markers[15] and conjunctions:

---

[12] It has been reported that pauses do not tend to coincide with syntactic boundaries (such as sentence, clause and phrase boundaries) in spontaneous speech to the same extent as they do in read speech, leading to the conclusion that the distribution of pause in spontaneous speech is freer (Garding 1967). It would be expected than, based on this, that the number of pauses occurring in places other than at IU boundaries was much greater in the present data.

[13] Disruptive utterances include false starts, repeats and filled pauses (Sabin *et al.* 1979).

[14] Notation: (narrative number : intonation unit number).

[15] Schiffrin (1987) defines discourse markers as sequentially dependent elements (words or phrases) which bracket units of talk, e.g. sentences or speech acts/dialogues moves.

e (0.78) e que de vez em quando passava uma coisa um/ laranja (05:07)
and *(0.78) and that once in a while an orange thing passed by*

aí o Mike falou "não não acorda ela pra dar (0.76) **né?** antibiótico" (07:18)
then Mike said "don't don't wake her up to give (0.76) ***you know?*** *antibiotics*

e (1.39) eu não sei se/ a gente jogava mui/ jogava muito na rua (13:03)
and (0.78) *I don't know if/ we played a lo/ played a lot in the streets*

(3) Accented content words:[16]

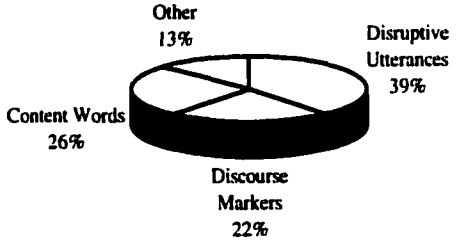
aquilo ali me deu (1.14) um **troço** tremendo (05:47)
that thing gave me (1.14) *an enormous* ***dread***

pra (0.98) **fazer** isso comigo (06:50)
to (0.98) ***do*** *that with me*

éh duas garotas viram um um ra/ um (0.85) uma (0.71) éh um (0.79) um **objeto** não-identificado (17:13)
ahn two girls saw a a ra/ a (0.85) *a* (0.71) *ahn a* (0.79) *an unknown* ***object***

The distribution of such pauses in terms of the linguistic phenomena associated with their occurrence can be found in Figure 2.2:

Figure 2.2
Distribution of the non-boundary pauses in terms of their association with different linguistic phenomena



[16] Content word analysis was based on traditional grammatical categories (Linsday & Cintra 1985).

As it can be verified, most of these pauses occur along with disruptive utterances (which are represented most of the time by the repetition of a specific word or clause or by the substitution of lexical items).

Most of these pauses also reflect a cognitive process, such as re-elaboration of a statement or idea, planning ahead and/or lexical choice. The following examples illustrate the fact that sometimes a pause that occurs before a content word can be also associated with the cognitive process of lexical decision-making:

> e entra a (0.64) entra a (1.24) a a *manageress* do restaurante
> a gerente (05:39-40)
>
> *and in come the* (0.64) *in come the* (1.24) *the the* manageress *of the restaurant*
> *the administrator*

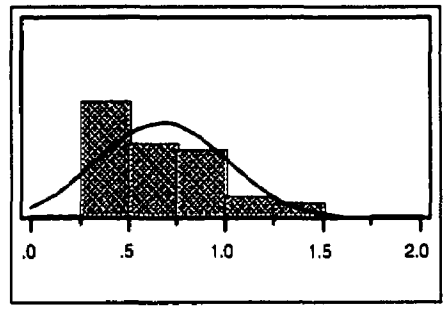
The narrator is telling a story that occurred to him while he was working in a restaurant in London. The repetition of the clause "entra a" suggests that he is searching for the appropriate word to use. His first choice is the English word "manageress," which would fit the context, not only because both the researcher and the participant were in an English speaking country at the time of the interview, but also because the narrative is situated in an English speaking country. Even so, he chooses to replace the word immediately after by the Portuguese equivalent: "gerente."

The distribution of pauses in terms of duration can be verified from Table 2.4. The mean duration of non-boundary pauses differs considerably from the mean duration of boundary pauses (0.68 sec versus 0.78 respectively). Although this difference is not statistically significant ($t=-1.577$, $df=337$, $p>0.05$)[17], a trend in this direction is easily established. This finding is supported by the literature (see, for example, Barik 1977;

Brown, Currie & Kenworthy 1980; Chafe 1989; Oliveira 1999; Rosenfield 1987; Sabin 1976, just to name a few). Pauses that do not coincide with a boundary are mostly very short in length, as Figure 2.3 below suggests:

Figure 2.3

Histogram of pause duration at non-boundary sites. Each bar shows the frequency of occurrence of the range of values represented on the axis. The line over the bar represents the normal curve

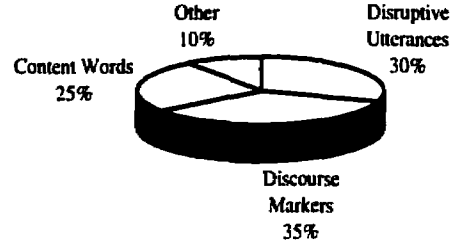


The histogram above (Figure 2.3) illustrates that most of the pauses at non-boundary sites are concentrated in the 0.25-0.5 sec. range. The distribution is quite skewed (Shapiro-Wilk $W=0.91$, $p<0.001$), as the normal curve suggests. There are just a few pauses above the mean pause duration (0.76). This finding is also in agreement with the literature (see Goldman-Eisler 1961; Goldman-Eisler 1972; Hawkins 1971; Maclay & Osgood 1959).

A few pauses that occur at non-boundary sites are rather long. An examination of the longer than average pauses that occur in this position reveals that they co-occur mostly with discourse markers:

[17] Note, however, that since the variances do not seem to be equal here (O'Brien's Test = 0.0049), a Welch Anova test was performed, allowing a significant effect ($F=3.83$; $df=1$ $p=0.0533$).

Figure 2.4
Distribution of the longer non-boundary pauses in terms of their association with different linguistic phenomena



Discourse markers generally reflect a cognitive process ("I mean," for instance, usually encodes a re-elaboration of a statement or idea) and for that matter a longer pause is likely to occur (Goldman-Eisler, 1961, 1972).

Long pauses sometimes may not be associated with any type of boundary, nor do they reflect a cognitive process. The next point to be discussed in this chapter has to do with a very common usage of pause in discourse, particularly in narratives: the exploitation of pause length for stylistic reasons.

2.8. Stylistic or semantic pauses

Longer pauses have been regarded in the literature as effective tools used to build tension within narratives. They are less common than the types of pauses considered so far, but functionally important. As Quirk & al. (1985) note, the speaker inserts a pause in discourse remarkably often for other purposes, such as before or after a particular word or string of words to obtain a certain effect. In doing so, "the speaker prosodically empathizes with the hearer" (Quirk & al. 1985: 1444). If the pause is longer than usual, its expressive, dramatic impact is evident.

Japanese storytelling is well known for the use of pauses with such particular function (the so-called "ma"). For instance, one of the most often used techniques of representation in Japanese storytelling consists of a simulation of a telephone-conversation, where the narrator pretends he is speaking with someone who is not present. The narrator provides a series of short confirmative or doubting phrases such as "yes-yes-who?-impossible..." and, in the meanwhile, makes pauses of considerable length. It is obviously used in order to maintain the tension of the audience about what is going on (see Balkenhol, Hollien & Hollien 1979 for more details).

Stylistic pauses are used with certain regularity in the narratives of the present study. They are closely related to the content of the narrative. For example, narrative 01 deals with an event of a very specific nature: it tells the story of a supernatural encounter. Many of the longest pauses in this narrative occur at climactic points in the story:

Narrative 01
Complication
13 — e e pai vinha por ali de noite (0.98) 14 — e ele começou a ver uma/ a a uh/ a sentir/ (0.43) ouvir passos atrás dele né? (0.9) 15 — e ele olhou pra trás 16 — olhou pros lado 17 — e não viu ninguém atrás dele 18 — ele co/ ai ele começou a correr assustado (1.26) 19 — e a/ e sentindo que a a coisa vinha corren/ continuando o se/ perseguindo ele né? 20 — vinha atrás dele correndo 21 — e ele correndo 22 — e a coisa correndo também (2.56)
Coda
23 — é uma experiência paranormal

Narrative 01
Complication
13 — *and my father was coming along in that night (0.98)* 14 — *and he started to see a/ to to uh/ to feel/ (0.43) to hear steps behind him (0.9)* 15 — *and he looked behind him* 16 — *he look at his sides* 17 — *and didn't see anybody* 18 — *he star/ then he started to run frightened (1.26)* 19 — *and to/ and to feel that the the thing was running/ keeping the fo/ following him you know?* 20 — *it was coming behind him* 21 — *and he was running* 22 — *and the thing was running too (2.56)*
Coda
23 — *it is a paranormal experience*

The evaluation section that precedes IU 13 elaborates on the deserted and spooky location where the narrative took place. The first clause in this story creates an atmosphere of suspense by informing the listener that the main character of the narrative

was passing by that spine-chilling place *during the night.* The pause that occurs at the end of IU 13 intensifies the suspense created by the information just given. It brings about the typical narrative question from the hearer: "what happened?" and, for that reason, generates involvement with the audience

The function of the long pause at the end of IU 14 is exactly the same as the one described above: it emphasizes the frightening character of the situation. A new aspect of suspense is introduced with: the sound of steps. The speaker's appeal to the element of mystery in this narrative is obvious. It is highlighted by the use of long pauses.

IU 18 reports the reaction of the main character to the enigmatic noise: once he realizes that there is nothing he could see that was making such a noise, he starts "*to run frightened.*" Once again, a long pause is employed. Now, more than ever, the suspense is at its highest point. This is the climax of the narrative. A long pause at this location is extremely relevant, as it reveals the "point" of the story. The long pause constitutes thus an evaluative feature in this narrative.

It is important to note that although all these pauses comprise some of the longest examples in this narrative, they do not occur at narrative boundaries: their function is primarily evaluative and not structural.

The longest pause in this narrative occurs at the end of IU 22. It has a twofold function: it indicates the end of both the complication section as well as the events of the story[18], and gives the audience an opportunity to digest the reported events and to eventually conclude what the evaluation implies and the coda makes explicit: that this is an example of paranormal experience. This particular example of the functions of pause

[18] The point of this story is to illustrate that paranormal experiences do happen to people that we know. Paranormal stories are characterized by mysterious, unexplained events. The lack of a resolution section in narrative 01 contributes to the atmosphere of mysteriousness typical of a narrative like this.

is discussed by Swerts & Geluykens (1994: 36) According to them, "the speaker pauses to give the listener the opportunity to take in the new referent cognitively, but also to enable him/her to reject the new referent if s/he should feel that way inclined."

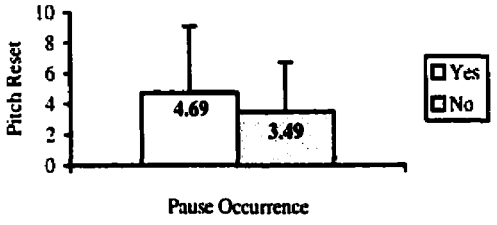
The distribution of pause in the corpus has been investigated. It is demonstrated that the hypotheses concerning the use of pause as a boundary predictor were confirmed: (a) narrative boundaries are marked by pause; (b) pauses that occur at narrative boundaries are longer than those that occur elsewhere. It is also found that pauses that do not coincide with any type of boundary have a very systematic distribution: they either can be associated with different linguistic phenomena or reflect a stylistic use. It seems then that semantic and pragmatic factors play a very important role in the distribution of pause in narratives. The question that may be raised at this point is whether pause occurrence and duration can be correlated with other prosodic variables.

2.9. Acoustic correlates of pause

Some recent studies have pointed out that pausal phenomena are closely related to acoustic features such as final lengthening, drop in intensity, vowel-like insertions, changes in voice quality and specific Fo patterns (see, for example, Hansson 1998; Hansson 1998, among others). In order to verify whether the other prosodic variables under investigation in this study are correlated with pause occurrence and duration, a series of t-tests and chi-square testes were done. For the purpose of this investigation, only boundary pauses were taken into consideration, as it would be impossible to account for some prosodic variables (such as pitch reset, for example) if non-boundary pauses were considered as well.

First we take a look at pause occurrence in relation to pitch reset.[19] Figure 2.5 below brings the mean and standard deviation of pitch reset as a function of pause occurrence:

Figure 2.5
Pitch reset as a function of pause occurrence



The numbers in Figure 2.5 clearly show that pause occurrence is associated with a higher pitch reset. The effect is significant (t=-3.870, df=608, p<0.0001).

Next, the relationship between pause occurrence and boundary tone is investigated.

Table 2.7
Correspondence of boundary tone with pause occurrence. The values are given in percentage

Boundary Tone	*Pause Occurrence*	
	Pause	No pause
Low	73	27
Non-low	61	39

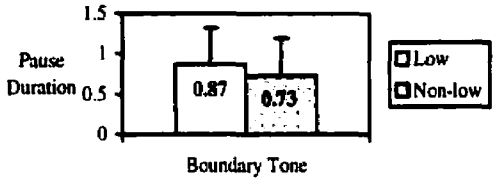
As the values in Table 2.7 suggest, pause occurrence is closely associated with low boundary tone (73% of the cases). The effect is also statistically significant (χ^2=50.749, df=609, p<0.0001).

[19] Pitch reset can be roughly defined as the difference of pitch range values of two adjacent intonation units. For a more detailed discussion on the notion of pitch reset, refer to Chapter 4, Section 4.2.2.

Finally, pause occurrence was correlated with speech rate reset,[20] but no statistically significant effect was found.

As for pause duration, significant effect was found only for boundary tone ($t=2.470$, $df=282$, $p<0.05$): longer pauses are associated with low boundary tone (see Figure 2.6 below).

Figure 2.6
Pause duration as a function of boundary tone



The results shown in Figure 2.6 clearly indicate that pause occurrence and duration can be predicted on the basis of several different acoustic factors. This does not mean that any of these prosodic variables determine the occurrence or duration of pause: these are statistically significant tendencies, not cases of absolute complementary distribution in the linguistic sense. The implications of the above-discussed correspondences for the present research are to be considered when all of the variables that were used for this analysis are individually studied.

It is important to note that the pattern is exactly the same for both narrative boundary pauses and non-narrative boundary pauses. In other words, a difference in

[20] Speech rate reset can be roughly defined as the difference of speech rate of two adjacent intonation units. See Chapter 3, Section 3.5 for a more detailed discussion on the subject.

terms of other acoustic features could not be verified in this data for pauses that occur at narrative boundaries and those that take place at non-narrative boundaries.[21]

In the following section, pause occurrence and duration are investigated as a function of individual narrative sections. If a correspondence between narrative sections and pausal phenomena is found, the hypotheses already tested and verified will achieve extra strength.

2.10. Pausing in narrative sections

Cognitive psychologists tell us that pausing reflects the strength or weakness of certain verbal habits. According to Goldman-Eisler (1968: 51), pauses are "synchronous with and indicative of encoding processes responsible for generation of information." In other words, pauses reflect the cognitive act of planning.

The complexity of information processing plays a central role in the use and duration of pause. In a series of studies, Goldman-Eisler (1961a, 1961b, 1968) compares the ratios of silent pauses in speech associated with tasks of varying levels of difficulty: the description of cartoons versus their interpretation. She found that the cartoon interpretations were correlated with longer pauses than the descriptions, which would support her initial hypothesis that, for a given task, the longer the delay between stimulus and response, the more cognitive operations are inferred as being required to produce the response. This finding has been replicated in many other studies (Siegman 1979;

[21] Note, however, that, in some cases, the statistical analyses fail to show a significant effect for the narrative boundary pauses, although a very clear trend could be verified. This may be due to the considerably small sample of narrative boundary pauses.

Reynolds & Paivio 1968; Levin, Silverman & Ford 1967; Lay & Paivio 1969; but see Rochester, Thurston & Rupp 1977, for a counterexample).

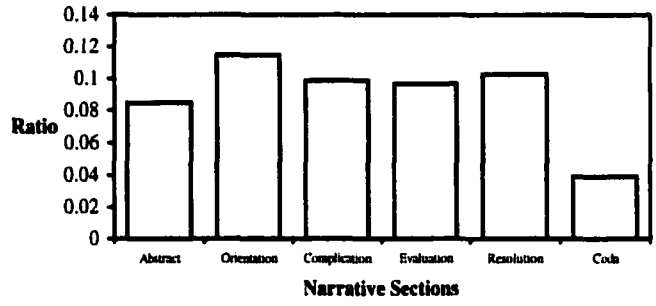
In a narrative, the sections that would be closely related to the cognitive process of interpreting would be the "evaluation" and the "coda," while the other sections ("abstract," "orientation," "complicating action" and resolution") involves the less complex process of "describing." As discussed in Chapter 1, "evaluations" are ways in which a narrator provides the point for telling her narrative. "Codas," on the other hand, signal the end of a narrative and are often used to return to the point of the conversation where the narrative was brought about. "Codas" not only function as a bridge to different speech modes, but, most of the time, also sum up the point of the narrative, often containing some sort of interpretation of the narrative as a whole. Since both sections are characterized by the more complex process of interpreting, one would expect the presence of a larger number of pauses of longer duration in such sections.[22]

In order to control for subject productivity, the concept of potential pause position (Brotherton 1979) was used for the analysis of pause occurrence in relation to narrative sections. Since a speaker can potentially pause between each pair of words in an utterance, a word boundary may be regarded as a potential pause position. In Figure 2.7 below, the number of pauses in each narrative section is expressed as a ratio of number of potential pause position in each section.

---

[22] In a study on pause strategy employed by native and non-native speakers of English in storytelling, Erbaugh (1996) observed that both groups of subjects paused far more often in passages of the narratives that would correspond to the Complication and the Resolution section in the Labovian terminology. According to her, this happens because of the importance of the information carried out in these sections: the more necessary the information is to the speech event, the more prone they are to pause.

Figure 2.7
Distribution of pauses as a ratio of potential pause positions in eacl of the six narrative sections



The numbers in Figure 2.7 contradict the above stated expectations. Neither evaluations nor codas are characterized by the presence of a larger number of pauses. As a matter of fact, codas display a considerably low quantity of pause, in relation to the other narrative section. While results of an analysis of variance showed statistically significant correlation between the occurrence of pause and narrative sections ($F(5,91)=2.419$, $p<0.0421$), post-hoc tests (Tukey-Kramer multiple comparisons) revealed that only codas differ significantly from all other sections.[23] It should be noted, though, that pauses occurring after the last intonation units in all the narratives were not counted, as they were not considered to be an integral part of the narratives. Since codas most often occur at the end of narratives, it follows that this might have somehow influenced the above results. Moreover, the analysis of variance showed a barely significant effect. It appears that pause occurrence was equally distributed in all narrative sections. Of course one could easily argue that this result is probably due to the questionably high cut-off point that was adopted here.[24]

[23] An analysis of the distribution of pause to speech ratio yielded a very similar pattern.
[24] Refer to the discussion in Section 2.2 above.

As for pause duration, no statistically significant effect was found that could substantiate the hypothesis that narrative sections are individualized by means of pause length. However, Table 2.7 below shows a trend in the direction of pauses being longer at the end of the narratives (in the resolution and coda), despite the high standard deviation in both cases. This could probably be interpreted as a means for the narrator to indicate that the narrative is approaching an end.

Table 2.7
Mean pause duration and standard deviation
for each of the six narrative sections

	Abstract	*Orientation*	*Complication*	*Evaluation*	*Resolution*	*Coda*
Mean	0.728	0.732	0.772	0.739	0.862	0.980
Sd.	0.343	0.369	0.405	0.460	0.615	0.806

In general, the above analyses refute the hypothesis that pause occurrence and duration varies as a function of the individual narrative sections. It seems that on the whole, the content of individual portions in a narrative has little or no influence on pausing strategies.[25] A closer look at individual narratives, however, will show that sections vary in terms of pause to speech ratio. If it is not the content of the narrative section that determines this variation, what would be?

2.11. The cognitive rhythm

Henderson, Goldman-Eisler & Skarbek (1966), Goldman-Eisler (1967), and Butterworth & Goldman-Eisler (1979) have argued that planning for units of discourse is

[25] Note that while this is true for all the narratives as a whole, individual cases should be considered: it is always possible that correlations based on a group of material may wash out important relations within cases taken individually. Some narratives (12 and 13, for example) show a pattern that at least in part confirm the hypothesis in question: pauses occur more often and are longer at evaluation sections and codas than anywhere else.

cyclical. They describe a pattern in which a period of long pauses and short speech bursts alternates with a period of little pausing and continuous speech. They call this pattern the "encoding cycle" and propose that it reflects an underlying "cognitive rhythm," in which the hesitant period would correspond to planning while the fluent phase would reflect the execution of this planning. As Rochester, Thurston & Rupp (1977: 62) accurately observe, "this portrayal of the role of pauses and the model of the speaker implied is attractive ... it suggests units of speech production which are functional rather than structural in nature, i.e., units based on cognitive activity rather than on linguistic analyses which may or may not be relevant to encoding processes" (see also Beattie 1984; Beattie 1980; Beattie 1983).

According to Butterworth & Goldman-Eisler (1979), what contributes to this variation are semantic factors, rather than syntactic ones. It is reasonable to expect then that this pattern would emerge in the material under investigation here. Since all narratives are segmented into sections that display a coherent semantic unit, one would anticipate the presence of an "encoding cycle."

Figures 2.8 and 2.9 below display the common pattern found in most narratives in the data. The cycle described by Goldman-Eisler and her associates is present here: a period of low pause to speech ratio is commonly followed by a period of higher pause to speech ratio in each narrative section.

Figure 2.8
Pause to speech ratio plotted over time for narrative 05

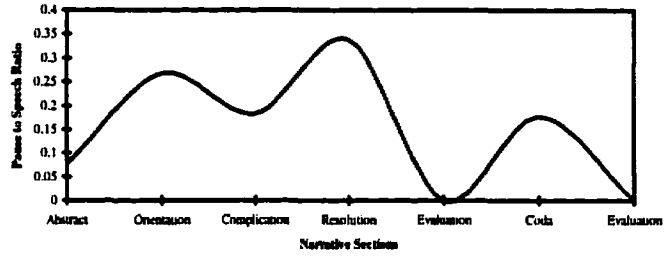
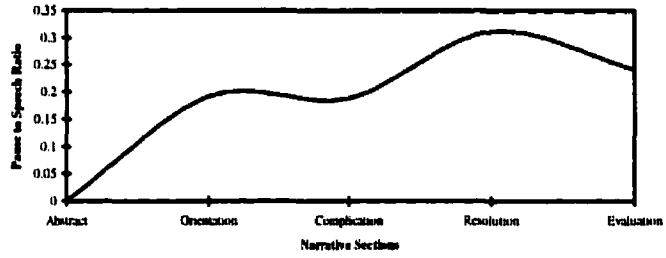


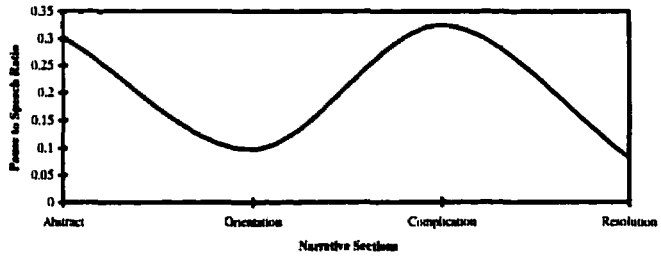
Figure 2.9
Pause to speech ratio plotted over time for narrative 07



Henderson, Goldman-Eisler & Skarbek (1966) and Butterworth & Goldman-Eisler (1979) propose that the "encoding cycle" is characteristic of spontaneous speech, as opposed to reading, which would probably not require the same amount of cognitive organization as non-prepared, spontaneous speech. Furthermore, Goldman-Eisler (1967) suggest that "cognitive rhythm" is more likely to occur in speech samples containing at least 30% of pausing and involving a cognitively demanding content. Although most of the narratives in the present study contain less than 30% of pausing (see Table 2.1 above) and bring no complex or novel content, the presence of the "encoding cycle" could be easily verified, even in narratives of very short duration, as Figure 2.10 below illustrates[26]:

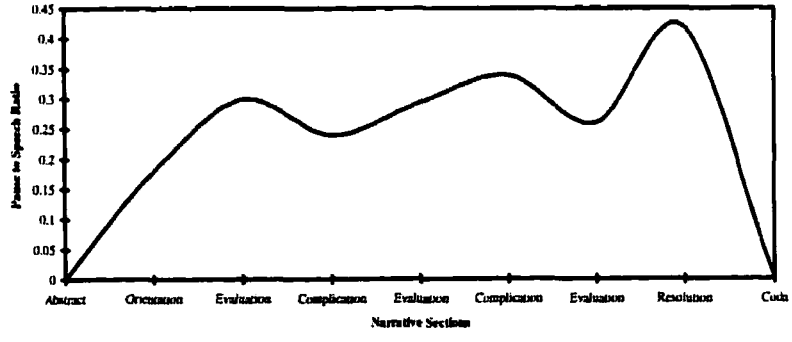
[26] Narrative 15 is the shortest narrative in the corpus, having the duration of 21.79 sec.

Figure 2.10
Pause to speech ratio plotted over time for narrative 15



Despite the fact that the "encoding cycle" can be characterized by the alternation of pause frequency in every other narrative section, as Figure 2.10 illustrates, this is not often the case. Some narratives do not display a difference in pause to speech ratio between narrative sections in some few instances. For example, in narrative 06 (Figure 2.11) the difference between pause to speech ratio in the orientation section and the first evaluation section is a positive one, i.e., the ratio in the first evaluation section is not lower than the ratio in the previous section, as would be expected. The same occurs in the boundary between the second evaluation section and the second complication section in the same narrative.

Figure 2.11
Pause to speech ratio plotted over time for narrative 06



It could be argued that this might be due to the degree of embedding of the narrative sections into each other. The more a narrative section is embedded to another in terms of its content, the less likely it will differ in terms of pause to speech ratio. Not surprisingly then, this phenomenon frequently involves evaluative sections: those sections have a higher degree of embedment with several other sections in a narrative, making it sometimes very hard to identify them as a separate unit. The distinction between the orientation section and the first evaluative section in narrative 6 is very loose:

Narrative 06
Complication
02 — eu me lembro quando eu era criança 03 — na escola
Evaluation
04 — menino como eu me lembro dessa cena 05 — poxa eu fiquei tão (0.43) triste 06 — tão triste no mundo (...)

Narrative 06
Complication
02 — *I remember when I was a child* 03 — *in school*
Evaluation
04 — *boy how I remember this scene* 05 — *I got so (0.43) sad* 06 — *so extremely sad (...)*

From a textual viewpoint, the orientation and the evaluation sections have a very close relationship: the evaluation elaborates on the information given in the orientation section. It even brings the same lexical item that is found in the orientation section (the verb "remember"). However, the information contained in the section that follows the orientation section in narrative 06 is primarily evaluative (the IU "*boy how I remember this scene*," for example, can be interpreted as both a *repetition* of the first IU in the orientation section, as well as an *intensifier* of the same IU).[27]

Since the first evaluation section in this narrative is so intrinsically related to the orientation section in terms of its content, the pattern described by Henderson, Goldman-Eisler & Skarbek (1966), Goldman-Eisler (1967), and Butterworth & Goldman-Eisler

[27] Repetitions and intensifiers are considered evaluative features by Labov (1972). See discussion of Chapter 1, Section 1.2.2.

(1979) fails to occur: there is no semantic dissimilarity between those sections that would justify the occurrence of the "encoding cycle" (Butterworth & Goldman-Eisler 1979).

The analysis of pause as a function of the content of individual narrative sections fails to show any significant effect. Nevertheless, the consideration of narrative sections as part of a cognitive cycle that would determine the frequency of pause usage resulted in a very favorable explanation for the observed variation of pause to speech ratio amongst narrative sections.

2.12. General discussion and conclusions

In this chapter, the role of pausing in narrative texts was discussed. It specifically dealt with the use of pause as a boundary indicator as well as an indicator of narrative section. The main questions are listed here again:

(1) Is pause occurrence predictable on the basis of information structure in spontaneous narratives? If so, how would pause occurrence and duration relate to narrative and non-narrative boundaries?

(2) Are different narrative sections characterized by variation in pause frequency?

(3) Is there any relationship between pause duration and different narrative sections?

It was hypothesized here, on the basis of the Labovian model of narrative analysis, that pausing reflects narrative structure — longer pauses are more likely to occur at the end of a narrative section than elsewhere in a narrative. Pause occurrence and duration would also be closely related to different narrative sections in a narrative text. Drawing from the findings of psycholinguistic research on pausological phenomena, it was also

hypothesized that narrative sections containing information that would require a higher amount of cognitive work, such as interpreting, as opposed to just reporting events, would present a greater number of longer pauses (and, consequently, a higher pause to speech ratio).

The analyses confirmed the hypotheses only in part. It was demonstrated that pausing is a strong indicator of narrative boundary: at the end of a narrative section, storytellers often produce a pause of longer duration to indicate that a given 'chunk of information' is completed and that a new 'chunk' is about to begin. This is in accordance with the findings in both psycholinguistic research (see Brotherton 1979; Butterworth 1980; Goldman-Eisler 1968; Maclay & Osgood 1959; Siegman & Feldestein 1979, among others) and computational linguistic research (see Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Hirschberg, Nakatani & Grosz 1995; Litman & Passonneau 1995; Litman & Passonneau 1993; Nakatani 1993; Nakatani & Hirschberg 1995; Passonneau & Litman 1997; Passonneau & Litman 1993, among others). Longer pauses are extremely important in the production of speech for two reasons. First, they give the speaker time to adequately formulate the next group of information. Second, they are very significant in speech perception, because they help the audience to cognitively digest the input.

No significant effect of pause occurrence and duration in relation to the various sections in a narrative was found. It was expected that sections containing any sort of interpretation, such as "evaluations" and "codas" would present a higher number of longer pauses than sections containing only descriptive information, such as "orientations" and "complicating actions." Since information dealing with the more cognitively complex task of interpreting requires a longer span of time to be processed,

the occurrence of longer pauses – which would be interpreted as the delay between stimulus and response in speech – would consequently be expected. However, the analysis did not show any relationship between narrative section and pausing strategies.

In several studies on the function and use of pause in discourse, Goldman-Eisler and her associates (1961a, 1961b, 1968) demonstrated that pause to speech ratio varies as a function of task complexity: people tend to pause more often and to remain in silence for much longer when performing the more complex task of interpreting the actions of a cartoon, rather than simply describing the same actions. It should be stressed that unlike Goldman-Eisler's studies, this work deals with non-elicited material. The fact that the subjects in Goldman-Eisler's studies were not familiar with the content of the material under investigation may have a decisive influence on the results there. The narratives in the present study are possibly part of each individual's repertoire of stories. Therefore we may expect less hesitation in the material used in this study.

Nevertheless, it was verified that a pattern displaying a period of hesitation followed by a period of fluency — which would correspond to the cognitive processes of planning and execution respectively — do exist in all narratives in the data and that this pattern reflects the way the narratives are structured: the periods of varying hesitancy generally correspond to the different sections in a narrative. It seems then that pausing phenomena (pause occurrence and duration) rather than being determined by the content of the information being carried by the individual narrative sections, is primarily governed by the cognitive rhythm of speech as proposed by Henderson, Goldman-Eisler & Skarbek (1966), Goldman-Eisler (1967), and Butterworth & Goldman-Eisler (1979). The "encoding cycle" was also verified by taking into consideration pause length within narrative sections and at narrative boundaries: in storytelling, speech seems to be

relatively hesitant at narrative boundaries (shown by longer duration of pause at these sites) and relatively fluent within narrative sections (as a consequence of the shorter duration of pauses within those units). The "Period of Planning" in this general semantic Plan (storytelling) would correspond to the narrative boundaries, while the period in which the Plan is executed would correspond to the narrative sections themselves. This would once more lend support to the concept of the cognitive rhythm.

A few other considerations were also made in this chapter concerning the use of pause in spontaneous, non-elicited narratives. On the one hand, it was verified that pauses occurring at locations other than boundaries (narrative and non-narrative boundaries) are very small in number. From a total of 338 pauses in the data, only 54 (16%) took place at non-boundary sites. The distribution of such pauses, rather than being haphazard, is quite predictable — almost 40% of the pauses that do not occur after an IU coincide with disruptive utterances (such as false starts, repeats and filled pauses); the other 60% co-occur with content words (26%), discourse markers (22%) and other elements (13%). With respect to duration, it was shown that the distribution of non-boundary pauses is highly non-normal: although the wide range variation (min. 0.21s, max. 1.76s), most of these pauses are located in the 0.25-0.5 sec. range. Longer pauses are most often associated with discourse markers (35%) and disruptive utterances (30%). These results are in agreement with previous studies on the same topic.

On the other hand, the occurrence and function of long pauses in narratives was also investigated. Pauses longer than 800 ms often have a stylistic function: most of the time, they serve to build the tension of a story. It was demonstrated that such pauses do not necessarily coincide with a narrative boundary, but that when they do so, they tend to be longer than usual. What we can conclude from this is that the content of a narrative may

have a great influence on the use and duration of pause. This finding also confirms one of the main hypotheses of this study, specifically that the content of the narrative may have an effect on the use and distribution of these prosodic variables (Chapter 1, Section 1.4.2, Hypothesis 4).

Additionally, it was observed that pause occurrence and duration may be associated with several other acoustic factors. Pause occurrence is closely associated with higher pitch reset and low boundary tone; while longer pauses co-occur with low boundary tone. These acoustic features and the implications that such correspondences have to the present study will be discussed in detail in the subsequent chapters.

This chapter investigated the use of pause as a means of revealing the structure of narrative discourse. It was found that pause is systematically employed to display larger "chunks of information" (narrative sections) – rather than of smaller, prosodically defined units (intonation units) – in this particular type of text. Pause duration also plays a significant role in this segmentation process: while pauses of shorter duration generally indicate that what follows is semantically and pragmatically related to the preceding information, longer pauses are often associated with structural breaks. Pausal phenomena, however, were not associated with the characterization of individual narrative sections. The hypotheses concerning pause usage in spontaneous, non-elicited narrative sections were then partially confirmed.

Pausing strategies are only one facet of prosody in the depiction of narrative discourse structure. While it was proved that this feature is rather reliable, it alone does not account for the manifestation of the structure of narrative texts. The next chapter will investigate one of two other features: the role of speech rate in narrative segmentation.

Chapter 3: Speech Rate

Be swift to hear, slow to speak.

James (1:19),
The New Testament

Abstract

Speech rate is examined in this chapter as an additional prosodic-temporal feature used to indicate the structure of narrative texts. The chapter starts with a brief review on the study of speech rate, followed by a discussion focusing on how it has been treated in the literature as segmentation cue in narrative discourse. Some methodological issues are discussed and the methods employed in the analysis are presented. Data regarding speech rate is then analyzed. The results suggest that the variation in speech rate that occurs within storytelling effectively contributes in marking out narrative structure.

3.1. Introduction

The study of speech rate is closely related to pausological research in many ways (see Chapter 2). 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.[1]

As the previous chapter demonstrated, 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.

The purpose of this chapter is to investigate whether speech rate variation is used as a cue for narrative segmentation. Rate will be analyzed on both the global and the local levels.[2] On the 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 found in the previous chapter, 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.

---

[1] This characteristic will be examined in more detail in Sections 3.2.3 and 2.8 below.
[2] For a definition of these concepts, see Chapter 1, Section 1.4.

In addition to relating speech rate variation to narrative segmentation, this chapter also addresses the role of rate as an evaluative device. It was found in the previous chapter that long pauses are often employed in narratives as an expressive tool. They are called "stylistic pauses," and are generally much longer than the speaker's average pause length, being evaluative in nature.[3] Based on this finding, it is expected that the use of speech rates that deviate from the norm have a function that is primarily evaluative.

The chapter is organized as follows: this introductory section commences with a brief chronological overview of the study on speech rate (Section 3.1.1), and later on focuses on the discussion of speech rate phenomena as a cue for narrative segmentation (Section 3.1.2). Section 3.2 addresses some methodological issues in the study of speech rate. This section is divided into three subsections, each of them discussing a particular methodological problem. The first is the multidetermination in the study of speech rate and the consequences of not taking this into consideration (Section 3.2.1). The second section addresses the units of measurement that are most commonly used in this type of investigation (Section 3.2.2). The third section looks into the distinction between speech rate and articulation rate and addresses the reason why only speech rate is adopted in the present investigation (Section 3.2.3).

Section 3.3 presents an overview of the speech rate observed in the data: general characteristics of all the narratives in the corpus are taken into account (Section 3.3.1), as well as individual characteristics of the participants in this project (Section 3.3.2). The actual analysis of the data starts in Section 3.4, which investigates the occurrence of a rate cycle in the narratives, suggesting that the manipulation of rate on the global level is a way of making the structure of narrative texts transparent. Section 3.5 examines the role

[3] Evaluation, according to Labov (1972), is everything that is unexpected, that deviates from the norm.

of speech rate on signaling narrative boundaries. The hypothesis concerning differences of rate between two adjacent IUs will be tested. The following section (3.6) explores the possibility of relating speech rate to narrative sections. The assumption is that some narrative sections will present a faster rate than others, due to the content that they convey. The analysis proceeds with the investigation of rate phenomena on the local level (Section 3.7). Here, intonation units are studied separately, in order to determine whether variation in rate is used as an evaluative device. Finally, some acoustic correlates of speech rate are examined in Section 3.8. Section 3.9 contains a summary of the main findings of the chapter together with some preliminary conclusions.

3.1.1. Early studies on speech rate: a general overview

One of the very first references to empirical research on speech rate dates back to Cattell (1886), who summarizes his findings as follows: "it takes about twice as long to read (aloud, as fast as possible) words which have no connexion as letters which make words... When a passage is read aloud at a normal rate, about the same time is taken for each word as when words having no connexion are read as fast as possible." Cattell's study dealt with several languages, such as English, French, German, Italian, Latin and Greek, as he was also concerned with the different rate employed by the same speaker while speaking a foreign language. According to him, the rate at which someone speaks a foreign language is determined by the familiarity that the person has with the language: the more familiar a person is with it, the faster the speech.

Beer (1910) was perhaps the first to introduce objective methodology for the study of rate in speech. Using a rather peculiar method of measurement that relied on soot marking from flames, he compared polysyllabic words with monosyllabic words, using the syllable per second unit of measure – a unit most widely employed today. In his experiment, he demonstrated that, for a given passage containing the same amount of syllables, polysyllabic words are read faster than monosyllabic words. He claims that this is due to the amount of meaning that is carried out in passages containing monosyllabic syllables: the larger the amount of meaning that is conveyed in a message, the longer the speaker will take to utter it.

Meaning is also considered to be a determining factor for the establishment of speech rate in Fröschels' (1920) study. He compared the repetition of nonsense syllables with the production of syllables articulated within words and found that nonsense syllables are often produced at a slower rate than those syllables that are part of real words. He also noted that people tend to be affected by curiosity when reading a passage of unknown content, which would result in a deliberate acceleration of the speech as a result of this curiosity.

Olson & Koetzle (1936) introduced developmental considerations to the study of speech rate, by investigating the speech of kindergarten children. They found that while boys tend to speak less than girls, they do so in a faster rate. Fairbanks & Hoaglin (1940) considered other external factors. According to these authors, emotions such as anger, fear, and indifference are closely related to fast speech rate, whereas contempt and grief are associated to slow rate.

Essen (1949) confirmed the hypothesis that meaning has a decisive influence on speech rate, by demonstrating that the more meaning there is invested in an utterance, the slower the articulation rate is. According to him, this has also to do with the emotional state of the speaker: psychological tension – a result of the demanding task of interpreting new meanings – would be the most immediate reason for speech rate variation.

On the basis of the overview above it may be concluded that the early studies on speech rate attempted to account for the observed variation by taking into consideration both semantic and emotional aspects apparent in the speech act examined. Methodological issues related to this type of research will be discussed in the following section.

3.1.2. 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 & 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 (Uhmann (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 & Hirchberg (1992) and Hirschberg & 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.

None of the above-mentioned studies attempted to determine a correlation between speech rate and narrative segmentation by taking into consideration a model that accounts for the way narratives are structured into semantically individualized sections. The purpose of the present chapter is not only to investigate whether speech rate is also used as cue for narrative segmentation, but also to explore exactly how it does. Variation on speech rate will be considered a function of specific narrative sections in the Labovian model of narrative analysis. If a cycle of varying rate is found, it will be considered as support for the hypothesis according to which speech rate is used as a segmentation cue. This will be confirmed by demonstrating that individual narrative sections display a specific rate behavior. Furthermore, this chapter aims to investigate whether differences in speech rate ("speech rate ratio") have an effect on narrative segmentation on both the local and the global levels.

The next section addresses several methodological issues that are closely related to the study of speech rate. Further, a justification for methods to be employed in the present investigation will be given.

3.2. Methodological issues in the study of speech rate

This subsection has three objectives. First, it aims to present a discussion of the variety of factors influencing the rate of speech, with emphasis on keeping the empirical data as homogeneous as possible. In so doing, it also attempts to legitimate the consideration of individual variances. It is a well-known fact that speech rate varies from person to person, from time to time within the same person's speech and from situational

context to situational context. Some of the various factors that are considered to be decisive in the variation in speech rate will be discussed below in order to demonstrate that speech rate is the result of a multitude of different functions.

Second, this subsection also provides an examination of the variety of units of measurement that are employed in the research on speech rate, with the purpose of justifying the selection of the unit that will be employed in this study. The units range from sounds per unit of time (Heinitz 1921; Hegedüs 1957; Essen 1949; Fónagy & Magdics 1960), to words per unit of time (Brubaker 1972; Barik 1977; Grosjean & Deschamps 1975; Grosjean & Deschamps 1972; Scollon 1981), to syllables per unit of time (Abercrombie 1967; Grosjean & Deschamps 1972; Meinhold 1972) and finally to beats per unit of time (Scollon 1981). This plethora of units of measurements employed in the literature not only reflects a serious methodological flaw – as it will be demonstrated below, but also makes the essential task of comparing results among various studies impossible.

Finally, this subsection also examines the so-called distinction between speech rate and articulation rate. The problems involving the study of both measures will be discussed and the adoption of speech rate, rather than articulation rate, in the present study as a major temporal phenomenon in the establishment of narrative structure will be rationalized.

3.2.1. Multidetermination

The complex influence of emotion and social purpose was clearly defined in Henze's (1953) study on the use of nonsense material in speech rate research. He demonstrated that the employment of such data is inappropriate because it does not reflect the weight that psychological and social aspects have in determining the rate of speech. He also pointed out that speech style (reading versus spontaneous speech) has a great impact in speech rate as well.

Hegedüs (1957) was among the first to observe that different types of material make use of different speech rate strategies. In his study on Hungarian broadcasted speech samples, he found that sports reports present, in general, a higher rate of speech, in comparison to other types of material (see also Fónagy & Magdics 1960). Sabin (1976) also found a significant variation between three types of speech: reading, spontaneous narration, and retelling. According to the results of his investigation, speech rate is fastest for reading, intermediate for spontaneous narration, and slowest for retelling. This finding was replicated in Sabin *et al.* (1979), who also noted that poetry is read more slowly than prose. Grosjean (1980) compared the results obtained in two previous studies (Grosjean & Deschamps 1973 and Grosjean & Deschamps 1972) and found a significant difference between interviews and descriptions of cartoons: the lowest rate in the interviews was the same as the highest rate in the descriptions.

Scollon (1981) compared several speech types using words per unit of time as the measurement of what he called "density." He found a higher density in radio speech activities, such as programme announcements (3.43 words/sec), sports news (3.19

words/sec), and a Groucho Marx show (2.92 words/sec), and a lower density in a conversation between parents and their one-year-old child (1.1 words/sec). Using the data of several other studies on temporal phenomena, Kowal, Wiese & O'Connell (1983) demonstrated that speech rate is indeed a legitimate tool used to characterize speech types. They compared storytelling and interviewees' speech and noted that speech rate for these two speech types are quite different. Barik (1977) also compared the rate of different types of spoken discourses and found that storytelling is one of the least fluent, while delivering lectures is often characterized by higher speech rates.

Speech style is often cited as one of the most frequent factors determining the variation in speech rate. Szawara & O'Connell (1977) investigated the relationship between temporal variables, such as pause duration, pause time and speech rate and spontaneity in the delivery of sermons. They found that spontaneous sermons (those delivered in churches) often present a slower speech rate than prepared sermons (those broadcasted in radio stations). Kowal *et al.* (1975) compared the rate of speech in spontaneous storytelling to that in oral reading and observed a faster rate during reading than storytelling (see also Blaauw 1995, Crystal & House 1990, Howell 1991, Laan 1997, Levin, Schaffer & Snow 1982, Koopmans-van Beinum 1990, and Koopmans-van Beinum & Van Donzel 1996 for similar results).

In addition to speech style, the age of the speaker also determines the variation in speech rate, as reported by Mysak & Hanley (1958). In a comprehensive study involving individuals ranging from 30 to 90 years of age, they found a significant decline in speech rate as a function of aging (see also Shipp & Hollien 1972). Kowal, O'Connell & Sabin (1975) found that this variation is also present in the early ages: speech rate tends to

increase steadily with age from kindergarten up to high school sophomores, with four distinguished age clusters: (1) kindergarten, (2) second graders, (3) fourth and sixth graders and (4) eight graders, sophomores and seniors. Sabin (1976), on the other hand, found no significant effect distinguishing college age subjects from middle age subjects. The response measures in his research yielded an effect of significant magnitude to distinguish the college age and middle age samples as one adult group from the child and adolescent samples altogether, suggesting a broad developmental trend. Finally, the research by Sabin *et al.* (1979) demonstrated that speech rate is slowest for young children, accelerates during preadolescence to a maximum level in adolescence, and then decreases in adulthood. According to Sabin *et al.* (1979: 40), while the faster rate found in adolescence may be an indication of the "high activity level" of that group, its decline in adulthood can be interpreted as "the gradual development of reflective and rhetorical dimensions of adult speech."

Language may also be a source of variation in speech rate, as the works by Hanley, Snidecor & Ringel (1966) and Hanley & Snidecor (1967) suggest. According to Hanley, Snidecor & Ringel (1966), native Spanish, Japanese and American differ in terms of phonation/time ratio.[4] Hanley & Snidecor (1967) include Tagalog in the list. However, both works suffer from a serious methodological problem, as they themselves suggest: the use of short speech samples questions the validity of the results.

Johnson, O'Connell & Sabin (1979), compared speech rate in narratives told by native speakers of Spanish and English and found that the Spanish speakers often employ faster articulation and speech rate. They suggest that this variation may be a result of the

[4] Osser & Peng (1964), on the other hand, found no significant difference between the rate of speech of Japanese and American speakers.

different strategies employed by the two groups: while the native speakers of English often used filled pauses on several occasions in their stories, native speakers of Spanish preferred parenthetical remarks.

Steer (1974) investigated the role of emotion in the variation in speech rate and found that emotion interacts with sex: female speakers tended to change the rate of speech to a greater degree and did so more consistently than males when speaking under different emotional conditions.

Clemmer, O'Connell & Loui 1979 suggest that level of proficiency also influences the characterization of speech rate: in a comparison between the readings of a passage from St. Paul made by beginning and advanced drama students, those authors found that the advanced students consistently employ a faster speech rate than the students in the beginning level. Kowal *et al.* (1975), in a series of three experiments, demonstrated that type of proficiency is reflected in the time pattern of speech.

Black (1950) considered the physical space that individuals occupy as a possible influence on the variation in speech rate. He found that people tend to read faster in smaller rooms and that "live" rooms generally motivate a faster speech rate.

Furthermore, the works by Clemmer 1975, Kowal *et al.* 1975, O'Connell, Kowal & Hörmann 1969 and O'Connell & Kowal 1972 clearly demonstrate that the complexity of speech is reflected in the multidetermination of temporal variables. These authors have investigated several independent variables, such as: social variables (socioeconomic status, social-situational context, cross-language effects); developmental variables (proficiency in a second language, and age/educational level); structural variables (punctuation, stimulus format, and syntactic complexity); cognitive variables (speech

task, semantic variation, glossolalia, bilingualism, and schizophrenic speech); and the gender variable. They conclude that most of these variables bear, to greater or lesser extents an intimate relation to the temporal aspects of speech, including speech rate.

The above discussion leads to the conclusion that the determinants affecting the rate of speech are varied. It is seldom an easy task to establish exactly what independent variables are effective at a given time, because of the complex process encompassing the act of speaking. Furthermore, it is clear that the operation of a determinant does not exclude the operation of others. This is exactly what O'Connell & Kowal (1983: 272) refer to as "multidetermination." According to O'Connell & Kowal (1983), most of the studies of temporal phenomena have failed to take into consideration this very characteristic of speech. The list of variables presented above is not to be considered an exhaustive list of possible explanations for the variation in speech rate, but rather it serves as an indicator of the multidetermination characterizing this temporal aspect of language.

The following words by O'Connell & Kowal (1983: 273) effectively sum up the topic of independent variables relating to speech rate: "Common sense should have led us to suspect, of course, that, as with nearly all very complex higher human processes, some people would talk faster, some slower, some coherently, some incoherently, some correctly, some incorrectly – all under the same influence of an anxiety inducing situation. In other words, the other side of the coin of multidetermination is individual differences or personal style. People differ in their reactions to complex situational stimuli in sophisticated ways, and even the same subject reacts differently from occasion to occasion." What is certain however is that it is always desirable to keep empirical data as homogeneous as possible in order to avoid misleading statements about different

samples. For that reason, consideration on variances that occur either as a function of different speech samples (narratives, in the present study, or as a function of different individuals is always in order.

3.2.2. Units of measurements

As mentioned previously in this chapter, 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. The purpose of the following discussion is to provide a brief analysis of the most important units of measurements as found in the literature in order to validate the employing of the unit chosen for the present study.

Words per unit of time were systematically employed as the unit of measurement in the early analyses of speech rate (Brigance 1926; Cantril & Allport 1935; Cattell 1886; Lumley 1933; Olson & Koetzle 1936; Snidecor 1943), and researcher continued to use words up until recently (Brubaker 1972; Barik 1977; Grosjean & Deschamps 1975; Grosjean & Deschamps 1972; Scollon 1981). However, words differ greatly in length, which has some researchers suggesting that the results are meaningless. Furthermore, as Scollon (1981: 14) has pointed out, "this factor is the least satisfactory in making

comparisons across languages since languages vary widely in the typical number of syllables per word as well as in the amount of information encoded within single words."[5]

The sound per unit of time measure of speech rate is not as widely used as word per unit of time. Heinitz (1921), Hegedüs (1957), Essen (1949), and Fónagy & Magdics (1960) are among the few who employed this unit of measurement. The problem with this unit of measurement is that certain phonetic characteristics that are typical of faster speech, such as contraction and assimilation, are not taken into consideration (O'Connell & Kowal 1972).

Another parameter for the temporal organization of speech has been suggested by Scollon (1981): beats per unit of time. Based on musical theory, he proposes that strong or accented syllables reflect the same rhythm that is found in, for example, music and poetry and that this rhythm would be the result of biological and ecological factors. According to him, the rhythm of speech is determined by the heart rate, which may be, in its turn, determined by the social context. Scollon (1981: 25) thus suggests, as a matter of conclusion that "we may literally be speaking our hearts."

Nevertheless, Scollon (1981: 19) proposes that the use of beats per unit of time (which would be the measure of tempo) should be regarded as an independent parameter of rhythm, not correlated with the perception of speech rate.[6] According to him, "the intuitive sense of rapid speech seems to be closely tied to density and not at all related to tempo." As a matter of illustration, he noted that while the analysis of a Groucho Marx show exhibited a relatively slow tempo (75.9 beats per minute), it displayed a

---

[5] From this passage, it is clear to see once more that semantic factors play an important role in the characterization of this temporal phenomenon of speech.

[6] See, however, Auer (1990) and Couper-Kuhlen (1992), for counterarguments.

considerably high density (2.92 words per second), resulting in giving the impression of high speech rate.

The best-suited unit of measurement for speech rate, according to many authors, is the syllables per unit of time (Uhmann 1992, O'Connell & Kowal 1972). Abercrombie (1967: 96), for example, defines speech rate as the "rate of syllable succession." This is the unit adopted by Blaauw (1995), Fon (1999), Grosjean & Deschamps (1972), Grosjean & Deschamps 1973, Grosz & Hirschberg (1992), Uhmann (1992), Goldman-Eisler (1961), Hirschberg & Grosz (1992), Van Donzel 1999, Wood (1975), to name a few. Even though, as Uhmann (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.

Once again 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.

3.2.3. Speech rate versus articulation rate

According to Wood (1975: 100), "it is customary to make a fundamental distinction between gross rates based on the total time of speaking (i.e. including pauses)

and net rates based on the period of actual utterance (i.e. excluding pauses)." There is a great diversity on the nomenclature that is used to distinguish both phenomena. Goldman-Eisler (1961), for example, calls "talking" and "articulation" rates what Wood refers to as "gross" and "net" rates respectively. Generally, this distinction is made in terms of "speech rate" (for "gross rate" in Wood's terminology or "speaking rate" in Goldman-Eisler's) and "articulation rate" (for what Wood names "net rate").

In order to justify the necessity of making a distinction between those two rate phenomena, consider the following passage extracted from narrative 01 of the present corpus:

> e e pai vinha por ali de noite (0.98) e ele começou a ver uma/ a a uh/ a sentir/ (0.43) ouvir passos atrás dele né? (0.9) e ele olhou pra trás olhou pros lado e não viu ninguém atrás dele ele co/ ai ele começou a correr assustado (1.26) e a/ e sentindo que a a coisa vinha corren/ continuando o se/ perseguindo ele né? vinha atrás dele correndo e ele correndo e a coisa correndo também (2.56) é uma experiência paranormal (01: 13-23)
>
> *and my father was coming along in that night (0.98) and he started to see a/ to to uh/ to feel/ (0.43) to hear steps behind him (0.9) and he looked behind him he look at his sides and didn't see anybody he started to run frightened (1.26) and to/ and to feel that the the thing was running/ keeping the fo/ following him you know? it was coming behind him and he was running and the thing was running too (2.56) it is a paranormal experience* (01: 13-23)

This passage contains a total of 138 syllables[7], uttered in 28.6 seconds, which gives a rate of 4.8 syllables per second. However, between the first and the last

[7] The counting was made excluding possible contractions, as to avoid subjectivity due to perceptual factors. The last intonation unit in this passage ("é uma experiência paranormal"), for example, contains exactly 12 syllables if the counting is made disregarding the actual speech sample. The examination of the speech material, however, reveals that apparently a contraction takes place between the words "uma" e "experiência." This would make the total number of syllables in this sentence drop to 11. The author of the present study asked a few colleagues to examine the speech sample, but there was no agreement among them concerning the existence of a contraction between the above-mentioned words. This would suggest that the perception of such phenomenon may vary, what would then make its consideration arguable. Of course that an analysis of the spectogram of the material would nullify any doubt, but this was found to be rather unimportant *for the present analysis*: the few number of cases verified in the whole data would not justify this time consuming and sometimes problematic task and would not likely cause any major influence over the final results. For example, in this entire passage, only three cases were considered to be candidates of some sort of contraction, what would result in only 3 syllables being disregarded from a total of 138. The influence over the final result would be extremely insignificant: less than 0.2 syllables! Furthermore, Uhmann (1992: 312) observes that "if speakers show syllable omission in passages which are

articulation, there are six pauses totaling 6.1 seconds. This means that this participant took only 22.5 seconds to produce 138 syllables, resulting in a rate of articulation of 6.1 syllables per second. It is clear from these numbers that we are dealing with two completely different concepts here. As Uhmann (1992: 305) puts it, speech rate "tells us how much time a speaker takes to communicate an idea," while articulation rate "tells us how fast s/he produces sounds." The questions to be asked are: what measure is the most significant in the perception of rate, and which of them is more relevant to the present investigation?

A number of scholars have pointed out that pause occurrence and duration have a great influence on the perception of speech rate. Goldman-Eisler (1961: 171), for example, clearly states that "a continuous flow of speech, rarely broken by periods of silence, is felt to be fast speech, and speech the flow of which is haltered by frequent pauses of hesitation is experienced as slow speech." (see also Grosjean & Deschamps 1972; Jespersen 1913 and Lass & Deem 1972).

Some authors, on the other hand, propose that it is not only the duration and frequency of pause that play an important role in the perception of speech rate: the position in which it occurs with respect to the segmentation of the speech material into intonation units has some influence in rate perception as well. Scollon (1982: 339), for example, argues that pauses occurring at the end of intonation units are qualitatively different from those that take place within intonation units. He suggests that only the pauses occurring at the end of intonation units have interactive and cognitive functions

---

perceived as 'fast' an extra factor must be taken into account depending on how many syllables are omitted in comparison to a standard or 'lento' pronunciation." According to her, "this factor contributes to the perceived 'fastness' to cover the fact that the same was said in a lesser amount of time." She concludes then that "just counting the number of actually pronounced syllables might miss this effect."

and refers to those pauses as "useful silences" (see also Hawkins 1971). Lass & Deem (1972), on the other hand, advocate the importance of pauses occurring within intonation units. After studying the speech of informants who were asked to read a given passage in normal, half as fast, and twice as fast speech rates, they observed an increase and a reduction of speech rate as a consequence of the changes in the number of pauses that are located within intonation units.[8]

Since it is not the fact of whether a speaker is articulating words at a given rate that will determine the perception of 'fast' or 'slow' speech, and since what is crucial is how the final product of speech will be perceived by the audience, one must conclude that pause should not be disregarded in the analysis of speech rate. Pauses, as demonstrated in the previous chapter, are extremely important in the communication of ideas[9] and discourse segmentation has mainly to do with to the way ideas are organized into a given frame.

For the reasons discussed above, speech rate and not articulation rate will be considered in the present investigation. This choice was made not only on the basis of the methodological grounds discussed above, but also because research on "articulation rate" has several drawbacks. For example, there are a number of authors who consider articulation rate to be the rate of speech calculated disregarding only the length of unfilled pauses (Goldman-Eisler 1968; Grosjean 1980; Grosjean & Collins 1979; Grosjean & Deschamps 1975; Grosjean & Deschamps 1972; Grosjean & Deschamps 1973; Kowal & O'Connell 1987; Sabin 1976; Van Donzel 1999). But what about the

---

[8] Rietveld & Gussenhoven (1987) demonstrate that intonation may also have an effect on the perceived tempo of utterances.

[9] The pauses that are present in the excerpt of narrative 01 analyzed above have a very significant semantic weight in the narrative, as discussed in Chapter 2, Section 2.8.

filled pauses, repetitions, false starts and other phenomena the like? Should they be counted as part integrant of the speech?

The measurement of speech rate was made by examining the waveform, on the speech-editing program 'Praat' (Boersma 1997). Pauses and nonlinguistic utterances were treated as individual units (syllables) and were included in the calculation of rate,[10] since, as Fon (1999:663) puts it, "they might be indicators of conceptual planning and their existence might also contribute to rate perception."[11]

The next section will present the general results of the measurements of speech rate in the entire data.

3.3. Speech rate in the data: an overview

The results that follow are based on two notions: the analysis of the data as a function of the individual narratives, and as a function of the participants that produced these narratives. Analyzing both aspects separately is justified by the fact that speech rate variation occurs not only among speakers, but also happens within the speech of the same speaker, as stated above. Since the data contains different samples produced by the same speaker, it seems to be imperative to investigate the characteristics of speech rate from both of these perspectives.

---

[10] Note that the pauses that occur at the end of the intonation units were considered to be part of them.
[11] 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 opted for methodological reasons (Fon 1999).

3.3.1. General characteristics of the narratives

It was demonstrated in the previous chapter that the average duration of all the seventeen narratives in the data was around 66 (65.51) seconds in length. The mean syllabic production per narrative was 353 syllables, which resulted in an average rate of 5.5 syllables per second. The length of intonation units within narratives averaged 1.77 seconds. When evaluating the distribution of syllables within intonation units, it was determined that there is a mean value of ten syllables per intonation unit. The raw numbers are given in table 3.1 below:

Table 3.1
Overview of temporal characteristics in the data: total duration of narratives (pauses included), total number of syllables, average duration of intonation units (IUs), average number of syllables per IU, and total speech rate, broken down for narratives and mean values over narratives. The duration values are expressed in seconds and the rate values are given in syllables per second. Standard deviations are given in parentheses

Narrative Number	*Total Duration*	*Total Number of Syllables*	*Average Duration of IUs*	*Average Number of Syllables per IU*	*Total Rate*
01	48.54	233	2.11 (1.6)	10 (6)	4.8
02	31.49	224	1.37 (0.78)	10 (5)	7.1
03	54.23	374	1.29 (0.68)	9 (4)	6.9
04	47.16	284	1.59 (0.74)	9 (4)	6
05	109.44	548	2.06 (1.23)	10 (6)	5
06	113.78	511	2.03 (0.99)	9 (4)	4.5
07	63.59	427	1.77 (1.3)	12 (7)	6.7
08	45.6	318	1.43 (0.62)	10 (4)	6.9
09	52.69	241	1.59 (0.8)	7 (4)	4.5
10	150.72	836	1.75 (0.98)	10 (5)	5.5
11	66.39	393	1.44 (0.84)	9 (6)	5.9
12	49.01	264	1.88 (1.11)	10 (5)	5.3
13	81.49	339	2.54 (1.47)	11 (6)	4.1
14	62.87	326	1.69 (0.62)	9 (4)	5.1
15	21.79	116	1.55 (0.76)	8 (4)	5.3
16	49.38	302	1.59 (0.77)	10 (5)	6.1
17	65.57	258	2.42 (1.58)	10 (7)	3.9
Mean	65.51	353 (164)	1.77 (1.06)	10 (5)	5.5 (1)

The average rate of 5.5 syllables per second in the present data is above the average found in many other studies dealing with temporal aspects of storytelling. As

noted above, narratives often exhibit a relatively low rate when compared with other types of speech (Barik 1977; Sabin 1976; Sabin *et al.* 1979). The literature generally refers to a rate ranging from 2 to 4 syllables per second in narrative texts. Sabin (1976), for example, found a rate of 3.16 syllables per second in his study on rate phenomenon in adult narratives in English. The average rate reported in Van Donzel (1999) for her data from spontaneous narratives told in Dutch was 3.652 syllables per second. Barik (1977) compared the rate of storytelling in French and English and found a rate of 2.4 syllables per second in English and of 2.2 syllables per second in French.[12] Grosjean (1973, 1975) reached the same conclusion by comparing the rates of French and English, but found much lower values (1.96 and 1.85 syllables per second in French and English respectively).[13] Kowal & O'Connell (1987) studied the temporal aspects of narratives told during and after watching a movie and observed that the difference in rate between the two conditions was not significant (1.78 syllables per second and 3.62 syllables per second respectively). Fon (1999) found that narratives told by Mandarin speakers average a rate of 4.84 syllables per second. Kowal, Wiese & O'Connell (1983) reported a rate of 3.43 syllables per second in English narratives, while Van De Water *et al.* (1987) mention a rate of 2.92 syllables per second in their corpus.[14]

---

[12] The values are actually expressed in words per minute in the paper: 149.9 and 131.8 respectively.
[13] The values in both studies are also expressed in words per minute: 118 and 111 respectively.
[14] These varied results may be a consequence of the use of different methods for determining speech tempo. For example, a method that does not treat pauses as syllables may give lower numbers for syllables per second. The results are however also dependent on several factors, such as the speech style, the size of the speech sample, and the definition of pause in terms of duration. As noted above (see footnote 9), the results derived from a measurement that included pause and nonlinguistic utterances in the present corpus did not differ from the results derived from a measurement that did not include these elements. This may be a consequence of the relatively high cut-off point of pauses adopted here.

The higher rate in the data may reflect an idiosyncrasy of the Brazilian Portuguese language[15], but may also be an indicative of the higher spontaneity characteristic of non-elicited narratives. All the above-mentioned studies dealt with spontaneous (i.e. not read) narratives, but they were in one way or another elicited; the low speech rate found in the majority of these studies may be a consequence of that.[16]

Analysis of variance was carried out in order to find out whether a significant effect of speech rate as a function of the individual narratives is present in the data. Speech rate differs significantly for the narratives ($F(16,623)=7.4955$, $p<0.0001$).[17] Tukey-Kramer tests reveal that narratives 02, 03, 07 and 08 differ significantly from most of the other narratives in the data, as they display a considerably higher rate than the average. It is noteworthy that narratives 02, 03 and 08 were told by the same participant (05). Individual variation will be discussed below.

It is interesting to observe that the average number of syllables per intonation unit in the narratives of the data (10 syllables) is comparable with Sabin's (1976) study (9 syllables). In terms of the number of words, the average size of the IUs (5.5) diverges from Chafe's (1989) and Pawley & Syder's (1977) analyses for English (around 7 words per IU)[18], being closer to Pakosz & Flashner's (1988) investigation for Polish (around 5 words per IU). Analyses of variance were carried out with 'syllables per IU' and 'words

---

[15] To the author's knowledge, no study on speech rate in Brazilian Portuguese has been done.

[16] It is interesting to note, however, that spontaneous material is often characterized by a slower speech rate, if compared with read, non-spontaneous material (Crystal & House 1990, Kowal *et al.* 1975, Levin, Schaffer & Snow 1982, Koopmans-van Beinum 1990 and Koopmans-van Beinum & Van Donzel 1996). This is generally assumed to be a consequence of the need for planning time when speaking unprepared. A couple of experiments have demonstrated that speech rate, rather than being a function of speech style, is actually dependent on the specific experimental task (see, for example, Blaauw 1995, Baltiner *et al.* 1994 and Laan 1997).

[17] The rate of each individual IU was used as the source for this analysis.

[18] Van Donzel (1999) reports a rate of 7.3 words per IU in her study on prosodic aspects of information structure in Dutch.

per IU' as dependent variables and 'narrative number' as fixed factor to test whether the differences were significant. No effect was found.

There exists a significant variation among the values of mean IU duration displayed in Table 3.1 ($F(16,629)=4.02$, $p<0.0001$). Post-hoc tests (Tukey-Kramer HSD) show two homogeneous subsets of narratives (narratives within one group do not significantly differ from one another): (1) 01, 05, 06, 07, 10, 12, 14, 15, and 17, and (2) 02, 03, 04, 08, 09, 11, 13, and 16.

It is important to note that the amount of freedom to vary speech rate is a function of the length of IUs: the range for rate in short IUs (i.e. shorter than the average) varied from very slow (0.8 syllables per second) to very fast (19.2 syllables per second). As IUs grew longer (i.e., longer than the average), this range narrowed down (1 to 10.4 syllables per second).

3.3.2. General characteristics of the participants

Table 3.2 below displays the numbers resulting from the analysis of the same temporal aspects evaluated above for all the participants in this study.

Table 3.2
Overview of temporal characteristics in the data as a function of individual speakers: total duration of speech samples (pauses included), total number of syllables, average duration of intonation units (IUs), average number of syllables per IU, and total speech rate, broken down for participants and mean values over participants. The duration values are expressed in seconds and the rate values are given in syllables per second. Standard deviations are given in parentheses

Participant Number	*Total Duration*	*Total Number of Syllables*	*Average Duration of IUs*	*Average Number of Syllables per IU*	*Total Rate*
01	118.26	499	1.97 (1.28)	8 (5)	4.2
02	115.77	695	1.50 (0.81)	9 (5)	6
03	213.59	1162	1.74 (0.89)	9 (5)	5.3
04	70.8	380	1.77 (1.01)	10 (4)	5.3
05	178.48	1200	1.41 (0.70)	9 (4)	6.7
06	109.44	548	2.06 (1.23)	10 (6)	5
07	63.59	427	1.77 (1.30)	12 (7)	6.7
08	243.81	1083	2.19 (1.28)	10 (5)	4.5
Mean	139.22	749 (345)	1.77 (1.06)	10 (5)	5.5 (1)

Participants 05, 07 and 02 clearly stand out from the group in terms of use of speech rate. They employed a relatively faster speech rate, compared to the other five participants. Analysis of variance was carried out, resulting in significant effect ($F(7,16)=15.31$, $p>0.0002$). Post-hoc tests (Tukey-Kramer HSD) confirm that participants 02, 05 and 07 significantly differ from all the other participants with respect to speech rate.

According to Goldman-Eisler (1968: 25), articulation rate "is a personality constant of remarkable invariance." It was mentioned in the previous subsection that three of the four narratives that displayed the highest rate of speech in the corpus were told by the same participant (05). It is not surprisingly then that participant 05 (along with participant 07, who was responsible for the production of the fourth narrative exhibiting the highest rate) presents the highest speech rate in the data: this is a characteristic of this particular speaker. Nonetheless, it should be pointed out that variation within the same speaker might exist as a function of speech style, discourse type and varied experimental

conditions (Kowal & O'Connell 1987), as discussed above. Since the present study makes use of very homogeneous data, such an observation could not be made.[19]

It was demonstrated in the previous chapter that the mean duration of the computed pauses was longer for participants 08 and 01. The assumption that pause duration have a decisive influence over speech rate holds true in the present research as well: both participants 08 and 01 present slower rates in the corpus.[20]

As for average duration of IUs and average number of syllables per IUs, special mention should be made to participants 06 and 08, for the relatively long mean IU duration (2.06 and 2.19 seconds respectively) and to participant 07, for the relatively high number of syllables averaged per IUs (7 syllables). Analyses of variance were carried out with 'syllable per IU' and 'IU duration' as dependent variables and 'participant number' as fixed factor to test whether the differences were significant. Results show significant effect only for 'IU duration' ($F(7,626)=6.54$, $p<0.0001$). Post-hoc tests (Tukey-Kramer HSD) reveal that participant 08 differs significantly in her mean length of IU from participants 02, 03 and 05, while participant 06 differs only from participant 05, who presented the smallest average IU duration of the entire group.

---

[19] It is noteworthy though that narrative 04, which was also told by participant 05, exhibits a much lower rate in comparison, for example, with narrative 02 (6 versus 7.1 syllables per second respectively). Although one may argue that one syllable per second does not constitute much of a difference, it should be emphasized that the uniformity of the data in terms of variables that could influence the rate of speech (such as discourse type, experimental condition and speech style) would, at least theoretically, imply a nearly constant speech rate for the same speaker. That is obviously not the case here.

[20] It should be noted though that there exists no one-to-one relationship between pause duration and speech rate in this case: participant 08 made use of a considerably higher pause duration average if compared with participant 01 (1.07 versus 0.85 seconds respectively), but she does not present the slowest rate as a consequence of that (4.5 versus 4.2 syllables per second respectively).

3.4. 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 the previous chapter that a cycle of varying fluency (stated as a measure of pause to speech ratio) does exist in the narratives analyzed in this study, 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 & Goldman-Eisler (1979), but also emerges as a function of the way narratives are structured.[21] Since pause occurrence and duration often has a decisive influence on speech rate, it is expected that this "speech rate cycle" emerges in the narratives of this corpus as well.

In a study on speech rate as a reflection of variance and invariance in conceptual planning in storytelling, Fon (1999) 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" (p. 666). 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. 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

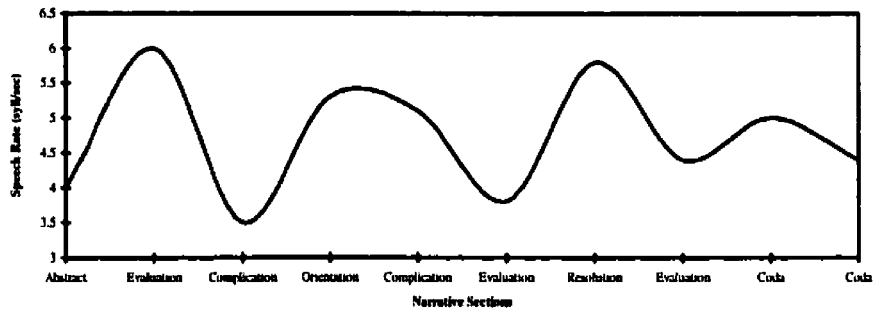
[21] Fitting the Labovian model of narrative analysis.

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.

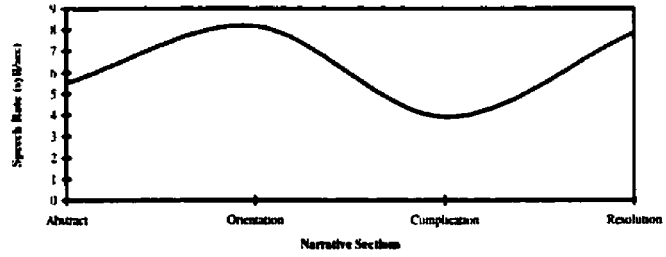
A typical rate cycle in the narratives of this investigation is given in Figure 3.1 below:

Figure 3.1
Speech rate plotted over time for narrative 14



The fluctuation of speech rate in this narrative, by a function of its integrant sections, is quite evident. There seems to exist a genuine 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 3.2 below):

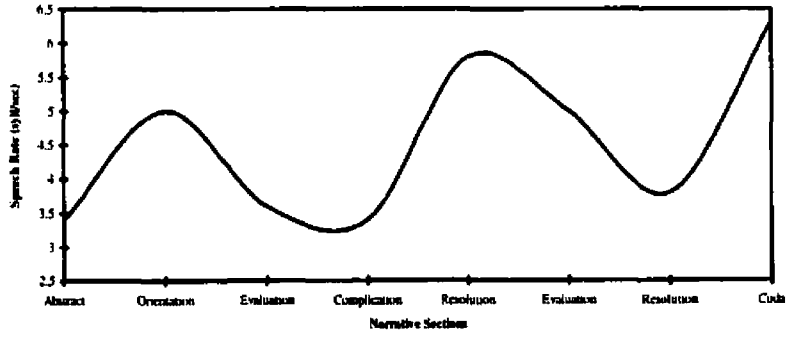
Figure 3.2
Speech Rate plotted over time for narrative 15



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 3.1), 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 3.2), 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 here again 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. It was verified that this is what explains the few exceptions in the case of the "cognitive cycle" analyzed in the previous chapter.[22] Since evaluative sections (or any other section that present 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 it is embedded. Narrative 17 (Figure 3.3) reiterates this point:

Figure 3.3
Speech rate plotted over time for narrative 17



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.

[22] It should be noted though that while in the case of the cycle of varying fluency this was considered to be an exception, here it is a norm (see Chapter 1, Section 2.11 for further details).

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 & Skarbek (1966), Goldman-Eisler (1967), and Butterworth & 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 reflects 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.

3.5. Speech rate reset

The previous chapter 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. 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 narrative sections.

3.6. Speech rate as a representation of narrative section

In an investigation of forms and functions of speech rate in conversation, Uhmann (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, Uhmann (1992: 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 Uhmann's working assumption for relevance is connected with the well-known distinction of given-new information.[23]

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.[24] 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

---

[23] This concept will be discussed in detail in Chapter 4.

[24] It would probably be a good idea to test subjective concepts such as "relevance", "speaker's involvement", and "dramatic scenes" (see discussion in the following sections for the use of some of these expressions) in terms of independent evidence. However, for the present research, which is mainly concerned with the employment of prosodic features in a specific theoretical model, such a procedure is not relevant.

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

Table 3.3
Mean speech rate values and standard deviations
for each of the six narrative sections (in syllables per second)

	Abstract	*Orientation*	*Complication*	*Evaluation*	*Resolution*	*Coda*
Mean	6.2	5.6	5.5	5.5	5.5	6.1
Sd.	2.2	1.1	1.2	1.8	1.8	1.6

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 narrative 08, for example, the orientation section 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:

Narrative 08
Orientation
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)
Orientation
07 - daí tava eu e meu outro irmão em casa (0.99)

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

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 & 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."[25]

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 of discourse that is

[25] See also Uhmann (1992), who found the rate of afterthoughts and summaries as topic-exit device to be faster than average in her study.

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. As discussed in Chapter 1, 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 discussed, 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.[26]

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

---

26 See, however, Koopmans-van Beinum & 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.

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 & 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, 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 (the numbers in the SR column correspond to the rate in syllables per second for each IU):

Narrative 05	**SR**
Orientation	
18- aí acho que em noventa e/	3.7
19- eu tinha catorze anos	10.4
20 - isso devia ser éh oitenta e quatro	7.7
21 - oitenta e cinco	7.4
22 - por aí (0.28)	3.3

Narrative 05	**SR**
Orientation	
18- *then I think that in ninety/*	3.7
19- *I was fourteen*	10.4
20 – *that must've been uh eighty four*	7.7
21 - *eighty five*	7.4
22 - *some time around this* (0.28)	3.3

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.[27] 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 parenthetic remark, or side comment, during the course of the story. Since both parentheses 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:

Narrative 09	**SR**
Complication	
08 – e áh aí eu acordei	4.2
09 – tava dormindo	7.5
10 – acordei (1.22)	1.7

Narrative 09	**SR**
Complication	
08 – *and uh then I woke up*	4.2
09 – *I was sleeping*	7.5
10 – *I woke up* (1.22)	1.7

[27] 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).

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

Narrative 16	SR
Evaluation	
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

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

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

IU 09, in the excerpt of narrative 09 given 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 is similar to the one extracted from narrative 09 only because the fastest IU communicates something that should have be 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 example, an excerpt 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.[28] 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 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.

3.7. Speech rate as an evaluative device

It was demonstrated in the previous chapter that pause duration can be manipulated in narratives as an expressive device. Longer pauses are often employed to build the tension of a given passage in a story, thereby being considered an evaluative strategy. This section aims to investigate whether speech rate can also be exploited as an evaluative feature.

---

[28] 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.

The first two examples illustrate the use of faster speech associated with dramatic scenes in narratives:

Narrative 01	**SR**
Complication	
13 - e e pai vinha por ali de noite (0.98)	3.1
14 - e ele começou a ver uma/ a a uh/ a sentir/ (0.43) ouvir passos atrás dele né? (0.9)	4.3
15 - e ele olhou pra trás	4.7
16 - olhou pros lado	9.4
17 - e não viu ninguém atrás dele	8.3
18 - ele co/ ai ele começou a correr assustado (1.26)	4.6
19 - e a/ e sentindo que a a coisa vinha corren/ continuando o se/ perseguindo ele né?	5.0
20 - vinha atrás dele correndo	8.3
21 - e ele correndo	12.8
22 - e a coisa correndo também (2.56)	2.5

Narrative 01	**SR**
Complication	
13 - *and my father was coming along in that night (0.98)*	3.1
14 - *and he started to see a/ to to uh/ to feel/ (0.43) to hear steps behind him (0.9)*	4.3
15 - *and he looked behind him*	4.7
16 - *he look at his sides*	9.4
17 - *and didn't see anybody*	8.3
18 - *he star/ then he started to run frightened (1.26)*	4.6
19 - *and to/ and to feel that the the thing was running/ keeping the fo/ following him you know?*	5.0
20 - *it was coming behind him*	8.3
21 - *and he was running*	12.8
22 - *and the thing was running too (2.56)*	2.5

Narrative 04	**SR**
Complication	
23 – a roda gran/ enorme	6.1
24 – começou a passar na frente do meu capô assim (1.13)	6.8
25 – daí a parte de trás do carro começou a subir (1.2)	6.0
26 – daí eu comecei a buzinar	12.5
27 – o cara parou	3.4

Narrative 04	**SR**
Complication	
23 – *the big/ enormous wheel*	6.1
24 – *started to come over the hood of my car like this* (1.13)	6.8
25 – *then the back of the car started to go up* (1.2)	6.0
26 – *then I started to honk the horn*	12.5
27 – *the man stopped*	3.4

The excerpt extracted from narrative 01 corresponds to the complication section of that story, where the events that the narrator considers to be "supernatural" are described. The higher values of speech rate associated with most IUs in this section can be interpreted as a result of the atmosphere of tension that the teller wants to create. The point of this narrative is to illustrate that "supernatural experiences" do indeed happen, that they are characterized by the mystery surrounding the events, and that panic often results. Fear is often associated with higher speech rate (Fairbanks & Hoaglin 1940).

Panic is shown to be the motivator for the faster rate of IU 26, in the above excerpt of narrative 04: the narrator is describing how a tractor almost collided with him

while he was driving his car. IU 26 is the climax of the narrative: it reports the fast reaction of the narrator that ultimately saved his life. Note that the use of the faster rate here not only represents the fear that the teller, as a character in the story, was feeling, but also functions as a mimicry of the action he is reporting. According to Uhmann (1992), that is precisely one of the functions of speech rate in discourse: to contextualize utterances in an onomatopoetic way.

The next example to be examined was extracted from narrative 08. This narrative tells the story of how the teller and his brother avenged their younger brother, who had been robbed by a homeless child. The excerpt describes the actions that followed when they found the very first homeless child on the street:

Narrative 08	SR
Complication	
20 - meu irmão tava tão bravo que quando ele viu um moleque	11.3
21 - ele falou p/ gritou pra mim "pára"	10.3
22 - eu parei (0.44)	3.6
23 - ele saiu do carro assim	7.7
24 - pegou o moleque pelo pescoço (0.75)	5.4
25 - falou "você assaltou meu irmão	10.1
26 - você não sei o que co/"	9.9
27 - batia no menino (0.56)	5.3
28 - nem sabia se era o via/ certo né? (1.35)	4.3

Narrative 08	SR
Complication	
20 - *my brother was so mad that when he saw a boy*	11.3
21 - *he said to/ screamed to me "stop"*	10.3
22 – *I stopped the car* (0.44)	3.6
23 - *he went out like this*	7.7
24 - *he got the guy by his neck* (0.75)	5.4
25 - *and said "you've stolen my brother*	10.1
26 - *you've/ I don't know"*	9.9
27 – *hitting the boy* (0.56)	5.3
28 - *he didn't even know if that was the right gu/ you know?* (1.35)	4.3

The events described in this narrative are of an unquestionable brutality: two mid-class adolescents harassing and ultimately slapping a defenseless street child, without even knowing whether that child was responsible for the crime of which their brother was a victim. Since the actions are considered to be socially unaccepted (as the evaluative comment on IU 28 suggests), a face-saving maneuver[29] is employed in order to minimize

[29] Goffman (1955) observes that everybody has a public self-image (face) that she/he wishes to protect and to have protected in an interaction. Brown & Levinson (1978) have further extended Goffman's notion of face by presenting the concepts of positive face, (i.e., the desire that others want the same thing that self

the negative image that the events convey. The teller makes a clear effort to posit himself as a rational person[30], while he suggests that his brother acted the way he did because he was moved by anger and lack of self-control. Speech rate seems to be used in order to make this opposition clear: while describing his brother's actions, the narrator uses a faster rate, which implies impulsive behavior. The narrator's action (IU 22) and comment (IU 28), on the other hand, are uttered in a slower rate, which indicates composure.

Slow speech is also associated with negative feelings, such as contempt, disappointment and grief (Fairbanks & Hoaglin 1940). The below example, extracted from narrative 06, illustrates the employment of slow speech as an indication of disappointment:

Narrative 06	SR
Complication	
21 – eu não me esqueço da cara do menino	6.2
22 – eu não me esqueço do nome dele	9.3
23 – Márcio o nome dele (0.53)	4.3
24 – não me esqueço (2.07)	1.9
25 – eu dei a caneta a ele	6.2
26 – o filho da mãe me deu um bombril (1.72)	2.8

Narrative 06	SR
Complication	
21 – *I don't forget that boy's face*	6.2
22 – *I don't forget his name*	9.3
23 – *Márcio was his name* (0.53)	4.3
24 – *I don't forget* (2.07)	1.9
25 – *I gave him the pen*	6.2
26 – *and the son of a bitch gave me a steel wool* (1.72)	2.8

Narrative 06 tells the story of how the narrator, when very young, made a real effort to buy a reasonably good present for her secret pal and got from him (who happened to be her secret pal as well) a steel wool (a jocular allusion to her face, which was full of freckles). The narrative is an illustration of the point stated in IU 01: "mankind is very despicable." In order to demonstrate how shocking that episode was to her, the teller says that she remembers everything, every minor detail associated with that incident, even the name of the secret pal. When she utters the name of the boy who had

---

wants) and negative face, (i.e., the desire that one's own needs and wants be unimpeded and unintruded upon). For a discussion on how the notion of face-saving can be applied to narrative texts, see Oliveira (1999).

[30] For a discussion on the concept of *positioning* in narrative discourse, see Bamberg (1997).

belittled her (IU 23), she does so using a considerably lower speech rate (if compared to the rate of the previous IU).[31] The following IU (24) is articulated in an even slower rate. It is noteworthy that IU 24 consists of a repetition of what was said previously (that she does not forget the details of the episode), therefore providing a piece of non-relevant information. Instead of using a faster rate, which would be expected in the case of non-relevant information, the message is uttered in a slow rate, as to emphasize the importance of the event to her. The despicability of the events is well exemplified in the contrast in terms of rate values between IUs 25 and 26: when the narrator tells what she gave to her secret pal, she does so using a normal rate; when she reveals what her secret pal gave her in return, she uses a slow speech rate, which indicates her negative judgment towards the boy's attitude.

All the above examples suggest that there is a strong relationship between information at the local level and speech rate. Storytellers manipulate the rate of their speech in order to evaluate the events reported in their narratives.

3.8. Acoustic correlates of speech rate

The purpose of this section is to investigate whether speech rate can be predicted on the basis of the other acoustic phenomena that are considered in this study.

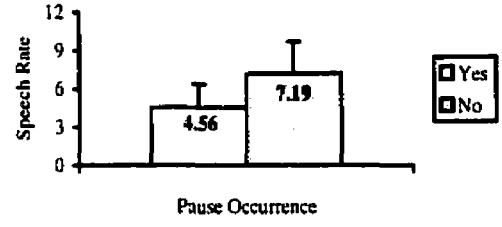
In the previous chapter, it was found that neither pause occurrence nor pause duration can be predicted on the basis of the size of speech rate reset. This clearly suggests that pausal phenomena are not related to differences in speech rate at the local level. However, a correlation between pausal phenomena and speech rate at the global

[31] It is important to note once again that sometimes it is the contrast among rates that contributes to the perception of fast/slow speech.

level was not clearly made. If it is true that one of the major determinants of speech rate is pause occurrence and pause duration, as the literature suggests (cf. 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 3.4
Speech rate as a function of pause occurrence



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 ($r=-0.52$, $N=627$, $p<0.0001$).

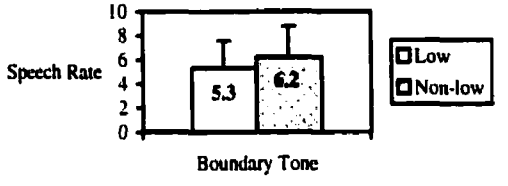
These results suggest that the occurrence of a pause within an IU guarantees that the speech rate of that IU will be higher that it would be if no pause was employed. On

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.

There is no attempt in the literature to correlate speech rate variation with other acoustic phenomena, such as pitch reset, pitch range and boundary tone[32]. In what follows is an attempt to identify a relationship between speech rate and the above-mentioned prosodic variables.

Speech rate values were correlated with pitch range, pitch reset, and boundary tone, but a statistically significant effect was only found for boundary tone. IUs ending in a non-low boundary tone are characterized by statistically significant higher speech rate values than those ending in a low boundary tone (t=-3.913, df=622, p>0.0001), as Figure 3.5 below reveals:

Figure 3.5
Speech rate as a function of boundary tone



This result may be interpreted in the light of the fact that non-low boundary tone is often associated with unfinished or disruptive utterances, such as those produced as self-repairs or parenthetical remarks (Brazil 1997). As discussed previously, these types of utterances are also associated with faster speech rate. Therefore, the result displayed in

[32] See Chapter 4 for a discussion of each of these concepts.

Figure 3.5 above validates the assumption that repairs and parentheses are often uttered in a faster speech.

3.9. General discussion and conclusions

The present chapter 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?

(4) What is the role of speech rate in narrative evaluation (if any)?

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

Based on the findings by Fon (1999), and on what was verified in the previous chapter with regards to the variation in pause to speech ratio in narratives (the "cognitive rhythm"), 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.

Rate reset seems to be closely associated with what Labov (1972) calls the "evaluative function" of narratives. It was observed that speakers often manipulate the rate of individual intonation units when they want to convey meanings that are not

available to the audience on the surface level of the message. This creates a rate contrast at the local level that serves mainly to clarify the point (or points) of the narrative. Rate reset then, rather than being used as a means of signaling the narrative segmentation into semantically independent sections, is employed as an evaluative device in storytelling.

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 possible relationship between rate and other prosodic phenomena was investigated. As expected, rate was found to be closely associated with pausal phenomena. Specifically, the occurrence of a pause within an intonation unit consistently occurs in periods in which there is a slower rate. Pause duration, however, did not have a significant effect over speech rate values. Of the few additional prosodic features that were taken into account in this investigation, only boundary tone interacted significantly with speech rate.

This chapter has dealt primarily with the role of speech rate in the signaling of narrative structure. In contrast to pausal phenomena, which plays a very important role in narrative segmentation on both the global and the local levels, 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. Contrary to what was expected, speech rate does not characterize narrative sections and is not manipulated on the local level in order to mark narrative boundaries. Manipulation of speech rate is only pursued at the local level as an evaluative device.

So far, only temporal aspects of speech were analyzed in this study on the role of prosodic features in narrative structuring. They both prove to be reliable indicators of narrative structure on both the local and global levels. The next chapter will investigate how pitch phenomena facilitate the job of making the structure of narrative texts transparent.

Chapter 4: Pitch Range, Pitch Reset, and Boundary Tone

Melody is a form of remembrance...
It must have a quality of inevitability in our ears.

Gian Carlo Menotti,
On patrons of the Metropolitan Opera, *Time* 1 May 50

Abstract

This chapter investigates the role of pitch in the organization of narrative discourse. It specifically deals with three intonational variables: pitch range, pitch range reset and boundary-marking pitch movements. The first variable, pitch range, will be examined on both the global and the local levels of narrative, in order to verify whether it is used as a means of characterizing individual narrative sections, and as an indicator of narrative section boundary. The other two variables, pitch reset and boundary marking pitch movements, will only be investigated on the local level as possible narrative boundary predictors. The chapter is structured similarly to the previous ones. First, a brief literature review will be provided. This will be followed by a discussion covering some methodological issues and a general overview of the data, in terms of pitch phenomena in relation to the (i) material analyzed, and (ii) the individuals who participated in this study. The analysis of the three variables was undertaken separately. A general discussion combining all aspects of the analysis is then presented. The main findings of the chapter support the hypothesis that these features play an important role in the display of narrative segmentation at different levels.

4.1. Introduction

Pitch, or the auditory correlate of the fundamental frequency of the vibration of the vocal folds[1], plays an important role in every level of speech: the segmental level (speech organized phonologically into syllables), the lexical level (speech organized morphologically into words), the syntactic level (speech organized into syntactic combinations), the topical level (speech organized into semantically determined topical constituents), and the interactional level (speech organized in an interaction by means of conversational turns).

At the segmental level, pitch helps, for instance, to distinguish between syllable-initial voiced and unvoiced consonants (Lehiste & Peterson 1961), and has an influence on the perception of syllabification. For example, a word such as *yes* can be perceived as bi-syllabic if it is produced by a falling pitch (Ainsworth 1986). At the lexical level, pitch can be used to indicate differences of lexical meaning or grammatical class, as in the noun/verb pair *ínsult* and *insúlt*.[2] At the syntactic level, pitch is often employed as a means of indicating major syntactic boundaries. For example, listeners often rely on pitch-contour variation, usually a preboundary fall-rise or rise, in the judgment of clause/sentence boundaries (Cooper & Sorensen 1977; Streeter 1978; Beach 1991; Wightman *et al.* 1992). At the topical level, pitch plays a very significant role in

[1] There is no agreement among researchers that the relationship between pitch and fundamental frequency is a clear one. Krause (1984: 243), for example, warns that "pitch perception is not to be equated with the perception of fundamental frequency of a periodic or quasiperiodic acoustic signal" (see also Pierrehumbert (1979) and Gussenhoven & Rietveld (1988) for a similar argument). However, some experimental studies have demonstrated that speakers usually compensate for nonlinguistic pitch perturbation, what makes the assumption that fundamental frequency is a reasonable indicator of pitch ('t Hart, Collier & Cohen 1990).
[2] Cutler & Clifton (1984) argue that stress information does not facilitate word recognition by English listeners.

determining accent, focus, and information structure. On the one hand, the assignment of new/given information in discourse is mainly relayed by means of fundamental frequency variation (Van Donzel 1999). On the other hand, discourse structure is often indicated through the use of fundamental frequency declination (Bruce 1982; Menn & Boyce 1982; Thorsen 1985; Ladd 1988). Finally, at the interactional – or conversational level, pitch, along with other prosodic cues, is used to construct and interpret turn regulating aspects of conversation, such as talking, keeping, yielding, and competing for the floor in a dialogue (Couper-Kuhlen & Selting 1996; Cutler & Pearson 1986; Selting 1992).

This study is concerned primarily with the topical use of pitch in discourse. The use of pitch at the segmental, lexical, syntactic and interactional levels will not be addressed here.

The present chapter is structured as follows: the introductory section provides a brief overview on the study of pitch phenomena in discourse (Section 4.1.1), a summary of the influencial works, and a discussion of the role that pitch plays in the structuring of discourse (Section 4.1.2) and specifically in narrative discourse (Section 4.1.3). Section 4.2 deals with some of the methodological issues of the investigation of pitch — specifically, the external variables that can affect pitch (Section 4.2.1), and the problems related to the measurements of pitch (Section 4.2.2). Section 4.2.2 also establishes the methods to be employed in the present analysis. Section 4.3 presents an overview of pitch behavior in the data. Most notably, general characteristics of all the narratives in the corpus are taken into account (Section 4.3.1), as well as individual characteristics of the participants in this project (Section 4.3.2). The analysis of the data will commence in Section 4.4, which investigates the occurrence of a linear pitch cycle in the narratives.

This will be followed by the description of a trend observed in the prosodic temporal variables previously investigated. The existence of such a cycle would suggest that speakers manipulate pitch on the global level as a means to make the structure of narrative texts more transparent. Section 4.5 examines the role of pitch range in the characterization of the different sections in a narrative. The assumption taken here is that some sections will present a wider pitch range than others, due to the content that they convey. The following section (Section 4.6) investigates the role of pitch in the signaling of narrative boundaries. In particular, it examines whether the difference in pitch range values between two adjacent intonation units are significantly higher in narrative boundaries in comparison with others elsewhere. The analysis proceeds with the investigation of boundary marking pitch movements as a cue for narrative segmentation (Section 4.7). Finally, some of the acoustic correlates of pitch are examined in Section 4.8, which concludes the analytical part of this chapter. Section 4.9 sums up the main findings of the chapter and provides some preliminary conclusions.

4.1.1. The study of pitch phenomena in discourse: a brief overview

There is a large number of studies published on the function of pitch in discourse. Far from intending to offer a comprehensive review of all available research, the aim of the present section is to provide a brief introductory overview of some of the representative research, focusing on a few key concepts, rather than on particular approaches.

One of the main functions that has been attributed to pitch in discourse is the role it plays in the complexity of the pragmatic organization of an utterance and its relation to discourse context. This has given rise to a number of works dealing with the binary categorization of information as 'new' or 'given' (Chafe 1976; Clark & Haviland 1977; Eady *et al.* 1986; Halliday 1967; Kuno 1972; Nooteboom & Kruyt 1987; Horne 1991; Fowler & Housum 1987). The basic idea behind these works is that an interaction can only proceed on the basis of the existence of a common ground between the speaker and the audience.[3] In the literature, the given/new distinction is connected to a number of related concepts, being presented under different names, such as old/new, theme/rheme, known/new, topic/comment, presupposition/focus, background/foreground, psychological subject/psychological predicate, presupposition/assertion, etc.[4]

Halliday (1963a, 1963b, 1963c, 1967a, 1967b, 1970) was perhaps the first to propose that information can be classified into what is *old, already known*, or *given*, and what is *new*. The old/new distinction is intrinsically a cognitive one, which could be defined as the difference between "that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance" and "what the speaker assumes he is introducing into the addressee's consciousness by what he says" (Chafe 1976: 30).[5]

The importance of the old/new distinction to the study of pitch phenomena lies in the fact that the ascription of what is old and what is new in an utterance is generally

[3] "Common ground is not restricted to shared experience of a particular linguistic interaction up to the moment of utterance; rather it is a product of the interpenetrating biographies of the participants, of which common involvement in a particular ongoing interaction constitutes only a part" (Couthard 1992: 41).
[4] See Lambrecht (1994) and Dryer (1996) for an overview.
[5] A detailed discussion of these concepts is beyond the scope of the present section. For general discussion from different perspectives, see Bolinger (1989), Chafe (1976), Culicover & Rochemont (1983), Haviland & Clark (1974), Ladd (1980), Prince (1981), Reinhart (1981), Rossi *et al.* (1981), Schaffer (1984), Selkirk (1984), and Van Donzel (1994).

made through the use of an *intonationally* marked focus (Halliday 1967).[6] The constituents of a message that convey communicatively important or new information is usually focused by the manipulation of pitch accents, while all the other constituents will generally not be as prominent (Van Donzel 1999).[7]

Furthermore, it has been proposed that speakers have a major choice between an end-rising *referring* tone, which is used to call on shared knowledge that has not been verbalized in the interaction yet, and an end-falling, *proclaiming* tone, which is employed to indicate the speaker's expectation that the area of common ground will be enlarged as a result of the speaker's being told something as yet unknown (Brazil 1997; Brazil, Coulthard & Johns 1980). The choice of tonic movement associated with the last syllable of an intonation group is referred to as *termination* (Brazil 1997). To clarify this notion, Brazil (1997: 68) presents the following example[8]:

// MARy BROWN (fall-rise termination) // is a TEACHer (fall termination) //
// MARy BROWN (fall termination) // is a TEACHer (fall-rise termination) //

The change of the termination tones in the sentence has a significant effect on the communicative values in these utterances, that are otherwise identical. The use of an end-rising, or *referring* tone, in both utterances, transmits the idea that the topic at hand is already in play in the talk, while the use of the end-falling, or *proclaiming* tone suggests that the conversational constituent was freshly introduced. In order to capture the difference of these utterances, Brazil (1997: 68) provides the following paraphrases of the two examples:

---

[6] *Focus* (or *emphasis*) may be defined as "the speaker's highlighting of part of an utterance" (Hirst & Cristo 1998: 31)

[7] It should be stressed though that pitch prominence is not the only acoustic feature that can be associated to focalization (or emphasis): intensity and duration may play a very important role as well.

[8] Upper case characters indicate prominent syllables; double slashes mark the key line; items that are in high or low key appear at the notational line.

'Talking of Mary Brown, she's a teacher.'
'Talking of teachers, Mary Brown's one.'

Besides indicating the termination tone of an utterance, a speaker must also select a relative pitch or *key* (Brazil 1997; Brazil, Coulthard & Johns 1980) for each and every intonation unit. Accordingly, *key* can be classified as a three-term system that is realized on the first syllable of the tonic segment – the onset syllable – and add the following meanings to the message: *contrast* (high key), *addition* (mid key), and *equality* (low key). In order to illustrate the different meanings attached to each of these keys, Brazil 1997(1997: 40) offers the following example:

// he GAMbled // and LOST // (high key and high termination)
LOST // (mid key and mid termination)
LOST // (low key and low termination)

The utterance that makes use of a high key indicates an opposition between 'gambled' and 'lost', which leads to the implication that 'he' usually wins when he gambles. The mid key choice suggests that he both gambled and lost. The use of a low key in the above utterance suggests that losing is a predictable consequence of gambling.

Notice that the above examples have only one prominent syllable in each intonation unit. In such a case, selection of a particular pitch level to realize a term in the key system necessarily involves selection of the same term in the termination system. Furthermore, it is noteworthy that, according to Brazil 1997(1997: 62), the difference between the key tone and the termination tone is never more than one 'level' in the three-term system. In other words, if in an intonation group the termination tone is high, the key tone will be either high or mid, but it will not be, under any circumstances, low.

Although a speaker has the option of realizing a proclaiming tone as either a 'fall' or a 'rise-fall', and a referring tone as either a 'rise' or a 'fall-rise', most often the choice

will be a 'falling' for the proclaiming tone and a 'fall-rise' for the referring tone. The 'rise-fall' and the 'rise' tones are thus a marked choice in the proclaiming/referring system. The use of the unmarked choice is restricted to certain situations. For example, the dominant speaker in a conversation may choose between the marked and unmarked choice, while the non-dominant speaker is restricted to the marked choice only (see Brazil (1997: 85ff) for further details).

Pitch movement has often been associated with various functions in discourse analysis. The meaning of an utterance is highly dependent on the type of pitch movement that occurs within the nucleus of an intonation unit. This particular feature is sometimes referred to as the *nuclear tone*. According to Cruttenden (1997: 50), the taxonomy of nuclear tones can be established upon three basic factors: (a) the initial movement from the nucleus, which can be *fall*, *rise*, or *level*; (b) the point where this initial movement starts in an utterance, which can be either *high* or *low*; and (c) an optional second change of pitch direction following the nucleus, which generates complex movements, such as rise-fall, fall-rise, rise-fall-rise, and so on.

Falling tones are often associated with a sense of finality, completeness, definiteness, and separateness, being therefore much more common within sentence final intonation groups. Cruttenden (1997: 91) proposes that the low-fall tone generally implies lack of interest, excitement and passion, whereas the high-fall carries the opposite meaning. The rise-fall tone, besides carrying the same sense of completeness that is characteristic of all the falling tones, also involves a sense of surprise, which is sometimes characteristic of exclamations grammatically marked as so, or, if accompanied by breathy voice, the effect of a gossip.

In contrast to falling tones, rising tones are often associated with non-finality and contrast. Cruttenden (1997: 94) suggests that the difference between low-rise, high-rise, and the mid-level[9] in non-final position is mainly a matter of style: "the low-rise is most oratorical and is also typical of a formal reading style, the high-rise is more casual, and the mid-level seems to carry no meaning other than that of non-finality, which is perhaps why it alone of these three tones occurs *only* in non-final position". In cases where the meanings of these tones do not depend on their relationship with the tone of another intonation unit, they may assume varied meanings. The low-rise is often associated with uncertainty when it occurs in declaratives. If it is not accompanied by a high pitch accent, an additional meaning of 'non-committal' or 'grumbling' is generally conveyed, so that this tone is quite similar to the low-fall. If it is accompanied by a high pitch accent, however, the meaning changes to 'soothing', 'reassuring' or 'patronizing'. The high-rise tone is generally employed in 'echo or repeat questions', which sometimes transmits incredulity. The fall-rise tone often carries a meaning of 'self-justification', 'appeal', and/or 'warning'. In declaratives, the fall-rise tone has often an element of 'contradiction' attached to it.[10] According to Brazil (1997: 87), what distinguishes the fall-rise tone from the simpler rise tone is that the latter carries a quality of 'peremptoriness', which is typical of a speaker in a dominant position.[11]

On the basis of the above discussion, it is clear that the research on pitch is closely associated with semantics and pragmatics, in the sense that it often attempts to relate intonational prominence and different pitch contours to convey meaning in specific

---

[9] Mid-level tone is grouped with the rising tones because it generally takes place in the same context, conveying similar meanings.

[10] For examples, see Cruttenden 1997(1997: 91ff).

[11] Note that in the present investigation, pitch movement will only be regarded at boundary sites. Those occurring at words bearing pitch accents will not be taken into account.

contexts (Halliday 1967; Bolinger 1972; Chafe 1974; Schmerling 1976; Gussenhoven 1983; Terken 1984; Baart 1987; Nooteboom & Kruyt 1987). However, as Pierrehumbert & Hirschberg (1990: 285) observed, most of these analyses have restricted the domain of interpretation to the phrase or utterance. Only in recent years have attempts been made to identify the "meaning" of intonational contours by taking into account the "meaning" of the whole discourse in which they occur.

In terms of levels of analysis, intonation can be approached in two different perspectives: (i) a phonological or 'abstract' one, which considers the acoustic parameter of intonation – fundamental frequency – as a cue for abstract categories used to model a formal system, and (ii) a phonetic or 'concrete' one, which takes the acoustic shapes of intonational contours and relates them to concrete meanings or functions (Ladd & Cutler 1983).[12] For a long time, the investigation of pitch behavior in discourse has been situated in an intermediate level of both approaches, with an emphasis on a holistic, or 'impressionistic' analysis. This could be justified by the limitations that research on intonation has imposed for an accurate analysis to a commonly large amount of material. Works that have such an impressionistic approach include, for example, that of Jones (1909), Palmer (1922), Armstrong & Ward (1926), Kingdon (1958), Quirk *et al.* (1968), Halliday (1967), Crystal (1975) and Svartvik & Quirk (1980), just to name a few. It was as late as the 1980s that research on pitch based on fundamental frequency data began to become widespread. The works by Pierrehumbert (1980), Cooper & Sorensen (1981), Willems (1982), Pijper (1983), and 't Hart, Collier & Cohen (1990) are just a few of the pioneering studies focussing on this relatively new approach. The present investigation,

[12] See, however, Hirst & Cristo (1998: 5), to whom "any attempt to define intonation on a physical basis... necessarily implies a formal (abstract) definition, even if this is never made explicit".

instead of adopting an impressionistic approach to the analysis of pitch phenomena, will concentrate, for the most part, on the physical realization of pitch in order to demonstrate (i) how fundamental frequency correlates with the organization of narrative structure, and (ii) what the relation of pitch is to the information that different narrative sections contain.

4.1.2. Pitch and discourse structure

A number of studies have been carried out concerning the role of pitch phenomena in the segmentation of discourse. Variation in fundamental frequency is generally considered to be a common device in discourse segmentation. More specifically, new topics are usually initiated with a relatively expanded pitch range and concluded with a compressed pitch range. Lehiste (1972, 1974, 1975a, 1975b, 1979, 1982), for example, found that high fundamental frequency is often associated with the beginning of discourse units. Utterances that are located at the beginning of such units are perceived with statistically significant accuracy as paragraph-initial sentences (see also Brown, Currie & Kenworthy 1980; Venditti & Swerts 1996; Swerts, Geluykens & Terken 1992; Swerts & Geluykens 1993; Swerts & Geluykens 1994; Nakajima & Allen 1992; Collier 1993; Garding 1982; Enkevist & Nordström 1978; Schegloff 1979; Brazil, Coulthard & Johns 1980; Butterworth 1975; Crystal 1969; Umeda 1982).

Sluijter & Terken (1993), on the other hand, found that in Dutch there is a tendency for fundamental frequency to decrease in the course of the paragraph in constructed texts. Nakajima & Allen (1992) observed that if there is no abrupt topic

shifting between utterances, fundamental frequency tends to be the same (see also Bruce 1982; Ladd 1988; Menn & Boyce 1982; Thorsen 1985; Hirschberg & Pierrehumbert 1986; Silverman 1987). Therefore, what seems to be an indication of topic change or discourse segmentation, rather than simply a high fundamental frequency of the first intonation unit in a paragraph, is the melodic discontinuity that occurs between two neighboring intonation units (Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Hirschberg, Nakatani & Grosz 1995; Swerts 1997; Hakoda & Sato 1980).

According to Hirschberg & Pierrehumbert (1986), systematic variation in pitch range can be used to signal a hierarchical segmentation of discourse: larger increases in pitch range mark major boundaries, whereas smaller increases indicate sub-segment boundaries. Therefore, while an increase in pitch range indicates discourse boundaries, a reduction in the amount of final lowering at a potential boundary can indicate that no such boundary in fact exists (see also Avesani & Vayra 1988 and Ayers 1992). Pitch range is also a very reliable predictor of labelers' consensus decisions on both global and local aspects of discourse structure. Grosz & Hirchberg (1992) demonstrated, for instance, that phrases initiating indirect quotations and parenthetical remarks could be distinguished from other phrases in almost 90% of the cases in terms of pitch range. Furthermore, Hirschberg & Pierrehumbert (1986) propose that pitch range can sometimes impose one segmentation of a discourse over another and can also disambiguate imprecise reference resolutions (see also Silverman 1987).

Pierrehumbert (1980) suggests that the relationship between utterances can be established by means of pitch accent, phrase accent, and boundary tone. Pitch accent (local fundamental frequency minima or maxima) is used to indicate the status of

discourse referents, modifiers, predicates, and associations specified by accented lexical items. Phrase accent (a simple high or low tone which leads the pitch movement from the last pitch accent of one intermediate phrase to the beginning of the next or the end of the utterance) indicates how an intermediate phrase relates to the surrounding ones. Boundary tone (which occurs at the end of the phrase, and can be classified as either high or low) is used to indicate whether a given phrase is "forward looking" or not.

The role boundary tones play in discourse as a signal of topic continuity or finality has been the focus of investigation in several studies dealing with the prosodic means of indicating discourse boundary (Blaauw 1995; Swerts, Collier & Terken 1994; Swerts, Geluykens & Terken 1992; Van Donzel 1999; Wichman, House & Rietveld 1997). Brown, Currie & Kenworthy (1980), for example, observed that low boundary tones are often associated with the end of a topic, while non-low boundary tones regularly suggest that there is more to come on the same topic.

Swerts & Geluykens (1994: 89) suggest that the binary distinction between 'continuation' and 'finality', as proposed by many studies, is "too crude". According to them, listeners generally identify the rounding off of a larger-scale discourse unit, based on melodic and temporal cues. They propose that prosodic features (especially boundary tones) can in fact indicate three discourse positions: "(a) 'there's much more to come', (b) 'there's one more to come' and (c) 'the description is completed'".

Using a more detailed approach, Swerts, Collier & Terken (1994) investigated the distribution of six different types of melodic contours in a route description: (a) early rise on last accent followed by continuation rise; (b) early rise on last accent followed by high declination; (c) fall on last accent followed by continuation rise; (d) early rise on last

accent followed by half-fall; (e) fall on last accent followed by low declination; and (f) late rise on last accent followed by late fall. They found a tendency for speakers to employ specific contour types according to the position of the utterance in the discourse. In this sense, contours of type (a), (b) and (c), mostly rising, were generally used in utterances that did not occur at final positions. Contours of type (d) characterized those utterances occurring in pre-final positions. Finally, contours of type (e) and (f), mostly falling, were predominantly reserved for utterances in final positions.

4.1.3. Pitch in narrative discourse

There exist only a few studies that have attempted to investigate the role of pitch phenomena in the organization of narrative discourse. Most, however, suffer from methodological problems that the present investigation strives to avoid. This section offers a brief summary of the main findings of some of these studies. Rather than attempting to cover a large number of works written on this specific topic, this section will concentrate on a few significant studies discussed in chronological order.

In a study on the use of pitch phenomena in oral narratives told in English by non-native speakers, Kumpf (1987) hypothesized that intonation units conveying evaluative information, or any other information that was considered to be backgrounded in relation to the events of the story itself would present a wider tone and pitch range variation. She observed that a wide pitch range characterizes irrealis situations, evaluative comments, complements, reported speech, contrasts, emphasized descriptions of a habit and continuous situation. High onset occurs much more often in non-event units, while low-

onset appears in afterthoughts and asides. A descending pattern was verified in descriptions of usual scenes. Based on these findings, the author concludes that pitch plays an important part in the structuring of narrative texts. The analysis presented in Kumpf's paper followed the iconic 'interlinear tonetic' system, which represents most of the time an auditory (impressionistic) analysis of the spoken material.

Pakosz & Flashner (1988) analyzed the role of prosodic features in Polish narratives and found, among other things, that these narratives were characterized by the presence of a rising-falling integration pattern forming a series of integrated units that usually terminates with a single falling tone unit. In order to establish whether these patterns would reflect the structure of the narrative, event line sequences, expository units, and foreground/background information were examined separately. Although expository units organized around one idea and the distinction between background and foreground information provided prosodic evidence for their status as separate units, no evidence was found that event line sequences in narratives show prosodic integration. The raising-falling integrated pattern found in Pakosz & Flashner's (1988) study could be easily interpreted as evidence of prosodically integrated narrative sections, although they did not relate these integrated units to semantically independent sections.

What seems to be worthy of mention is the fact that where event line sequences in the narratives analyzed by Pakosz & Flashner (1988) presented a prosodically integrated pattern, they did so by displaying the events in a "listing manner", with the last event in the series receiving a cadential interpretation. This "listing" effect could be interpreted in light of the fact that the narratives in this study were elicited.[13] Since narrators were

[13] The authors used the same method employed by the so-called Pear Story film group (Chafe 1980) to elicit their stories.

largely intended to describe the events that occurred in a movie they had just seen (*The Pear Story*), it seems perfectly acceptable that they did so using an intonation that resembles a "list". Narratives are told to make a point, and the point of elicited stories is, most of the time, to answer a request.[14] Pitch apparently facilitates the job of making the point of a story more transparent.[15] It would be of course desirable to verify whether such a pattern would occur in spontaneous, non-elicited narratives as well.

Pakosz & Flashner (1988) also observed in their narratives the occurrence of "extra falling tones" after summarizing comments, afterthoughts and elaborating remarks, which, in most cases, were accompanied by the use of monotone, narrow range and piano articulation. This fall would break the "rising-falling" integration pattern, which would be otherwise computed as such. If cases like this were included in the final computation of occurrences of the integration pattern, the percentage of the integrated units and subunits in their study would increase considerably. Finally, they also found that speakers sometimes make use of other prosodic features for signaling attenuation (elaboration, repetition, explanation, recapitulation, and recalled action) and highlighting (juxtaposition or contrast). Such features include monotone, low and narrow range, allegro, accelerando, and diminuendo articulation for attenuation, and wide pitch range, lento, and forte articulation for highlighting.[16]

Another significant study concerning the role of pitch phenomena in narratives is that of Selting (1992). In an attempt to investigate the function of intonation as a

---

[14] For a more detailed discussion of this particular matter, see Oliveira (1999).

[15] This assumption will be examined later in this chapter.

[16] It's important to note, though, as Pakosz & Flashner (1988: 44) did, that the speakers do not make use of such prosodic devices in a consistent manner: "there's no way of predicting that a given rhetorical effect will of necessity be marked with a certain prosodic device". Generalizations with respect of the use of prosodic features as a rhetorical effect should then be avoided (see Selting 1996: 238 for a similar argument).

contextualization device in the organization of storytelling, this author conducted an "auditive" analysis of narratives extracted from informal conversations recorded in laboratory conditions. All the narratives were told as an argumentation to support the narrator's point, making the data quite close to what one would find in a natural storytelling. In her analysis, Selting (1992) found that pitch was used both to differentiate between the storytelling activity and the conversation surrounding it, and to discriminate between individual parts within the story. In doing so, she observed in one of the narratives an almost exclusive use of falling accents in the conversational part surrounding the narrative. The narrative starts with a rising accent and develops with an alternation of falling and rising pitch movements, making the telling "livelier and more musical than the more serious sounding argumentation" (p. 243). Even more interestingly, she found that the first accent in an accent sequence generally follows the last accent movement of the preceding accent sequence, generating a link that displays cohesion: "accent types come in pairs, then they alter" (p. 243).

Pitch was also interpreted as a contextualization cue in Selting's (1992) analysis in that it helped establishing the difference between interpretative frames such as "same or different activity/subactivity", "more or less cohesive", "more or less expected", "more or less important or in the foreground", "more or less like recognizable activity types like lecturing, reading aloud, telling a fairy tale, chatting" and so on. The analysis of one of the narratives in her corpus demonstrated that intonation can be used to separate an entire evaluative section by means of pitch movement cohesion, corroborating the value of this prosodic feature in the segmentation of narrative discourse. Further to this, pitch was found to be employed as a rhetorical device in at least two occasions: (i) the building up

of the suspense in part of the narrative, generated through the use of systematic use of rising accents, and (ii) the signaling of lack of importance in the locutionary content, suggested by the employment of low pitch level.

The problem in Selting's (1992) analysis however, as Schwitalla (1992) observed, is that she made use of an impressionistic approach only. Since there was no attempt at any point to demonstrate that the transcription she offered could be validated from acoustic evidence, the whole analysis automatically incurs the risk of being considered biased and arguable. Schwitalla (1992: 260) himself, for example, who had access to the data analyzed by Selting, noted that he would identify or weigh accents and intonational contours differently.

According to Schwitalla (1992), the most important role pitch contours and repetition play in narrative discourse may be that of marking structurally different parts of a story. Using the Labovian evaluative model of narrative analysis, Schwitalla (1992) investigated how pitch phenomena helps to separate pieces of utterances in a narrative carrying communicatively different values. Based on the analysis of two parts of a single narrative, Schwitalla (1992) observed that global pitch contour and local pitch movement do indeed discriminate story parts, thus corroborating Selting's (1992) findings. However, he also noticed that rather than being marked merely by the switch of accent types, different illocutionary meanings are, for the most part, separated by "a combination of different prosodic features". Similarly to Selting's (1992) analysis, Schwitalla (1992) does not provide any acoustic evidence of his analysis of pitch phenomena in his data. So, the same criticism he makes on Selting's method also applies to his own. Furthermore, although he makes use of an independent model of narrative

analysis, which would, in theory, help to avoid the so-called risk of circularity[17], there is no attempt in the paper to demonstrate that the model he is using is replicable.

The works by Grosz & Hirchberg (1992) and Hirschberg & Grosz (1992) are examples of analyses that are supported by an acoustic examination of the material under investigation. In both works, these authors conducted corpus-based empirical work on intonational features using AP news stories recorded by a professional speaker. The main goal of the study was to examine the relationship between a number of acoustic-prosodic features and discourse structure, as determined by the independent model of discourse developed by Grosz & Sidner (1986). They found statistically significant correlations between prosodic-acoustic characteristics (pitch range, contour and pitch range change included) and both local and global aspects of discourse structure. Their findings are clearly not restricted to narrative texts, but are intended to subsume any type of discourse. However, instead of using spontaneous, non-elicited material, these authors opted for the examination of elicited narratives *read* by a professional newscaster. This fact obviously limits the extent of their findings.

Following his analysis on the communicative value of pitch in discourse, Brazil (1997: 92) proposes that the distinction between different types of rising (or *referring*) tones and falling (or *proclaiming*) tones is best exemplified in the act of telling stories. In parts of a narrative where the audience is being asked to recall an already shared tradition, the tone of choice is the marked referring tone (rise). If, however, the audience is made to feel that the information is new, an unmarked proclaiming tone (fall) will be used instead. Narratives then, would be divided into groups of intonation units characterized by the exclusive use of proclaiming tones (the reporting of events, for

[17] See Chapter 1 for further discussion of this matter.

example) and intermediate sections containing almost exclusively intonation units displaying referring tones (some evaluative comments, for example). According to Brazil (1997: 93), there is a predominance in narratives of a marked referring tone, and this could be explained if one takes into account the fact that in the act of storytelling, the narrator exercises control over the communicative event: "the narrator must not only establish his/her position *as* narrator at the beginning, but continuously reiterate his/her claim to dominant status – repeatedly insist on his/her expectation to be allowed to go on uninterrupted – until the end".

It can be clearly seen that most of these works give priority to a discourse analytic approach to deal with an issue that is primarily acoustic in nature. This is obviously not a problem *per se*. What might be objectionable, from a methodological standpoint, is the exclusive adoption of a tentative phonetic observation, based solely on subjective acoustic judgments that are not acoustically validated. If it is true that "it is necessary always to preserve a certain distance from the phonetic fact" if one wants to obtain clear and objective knowledge of the role of pitch in discourse, as Brazil (1997: 4) puts it[18], it is also true that "we can use intuition to *build theories*, i.e., to posit general underlying principles first, but their validity has to be tested by confirming that their predictions agree with the measurements" (Fujisaki 1997: 28). The present study will consider pitch phenomena as having both underlying principles and measurable manifestations that must be accounted for throughout an integrated process of deduction and induction. The

[18] "When we say that there is a 'fall' at a particular point in the utterance, it is not the fact that the pitch falls that we wish to make focal; it is rather the function or the language item which carries it, a function which, as far as we can tell from the examination of a lot of data, is typically realized by a falling pitch" (Brazil (1997: 4).

analysis in this work will therefore be guided by both acoustic and perceptual considerations of the data.

4.2. Methodological issues in the study of pitch phenomena

The current section will present a brief discussion of the variables that can affect fundamental frequency values, as well as introduce the methods to be employed in the measurement of the intonational features currently under investigation. Once again, the intent is to show that prosodic variables, as most linguistic manifestations, are the product of several different variables that should always be considered in any study that aims at providing general rules in a systematic way. It is also the purpose of this discussion to draw attention to the necessity of dealing with homogeneous and controlled data, in order to avoid generalizations over different types of material.

4.2.1. Some variables that can affect fundamental frequency values

Fundamental frequency is the product of several sources of variability, such as speaker attitude and emotion, segmental composition of the utterance, thematic accent placement, length of the linguistic domain of a contour, personal pathological features, personal physiological features, discourse condition, lexical and/or grammatical stress rules, etc. (Jassem & Demenko 1996). Although the consideration of such variables is, in most cases, imperative in any study that deals with the identification of pitch patterns, it is important to consider to what extent they actually interfere in the final product. The

present subsection will provide a short discussion of some of the most influential variables in the study of pitch phenomena with the sole purpose of endorsing the already stated argument that prosodic phenomena are the result of a number of variation sources that should be taken into account, if generalizations are to be made. The variables are presented and discussed below in a random order.

Emotion – or *attitude*, as some scholars prefer to name it (Couper-Kuhlen 1986; Uldall 1960)[19], is of central importance for the characterization of the overall pitch contour of an utterance. Uldall (1960 and 1964), for example, conducted a perceptual experiment that attempted to correlate affective meanings to intonation patterns. She found that the adjective 'pleasant' was often associated with rises ending high and a change of direction, while the adjective 'unpleasant' was characterized by a raised weak syllable and a narrow pitch range; 'authoritative' was associated with a wide range, a change of direction, a rising ending at mid and a final fall, while 'submissive' was characterized by a rise ending high. Fónagy (1978) also attempted to correlate emotion and pitch behavior. After examining the laryngographic recordings of ten texts passages read aloud, each exemplifying a given emotional state: (1) tenderness, (2) anger, (3) repressed anger, (4) joy, (5) fear, (6) sadness, (7) coquetry, (8) disdain, (9) longing, and (10) reproach, he concluded that "joy was characterized by a high pitch-level and large melodic intervals; a low average pitch-level and narrow intervals correlated with sorrow; fear was reflected in mid-high and reduced intervals" (p. 35).[20]

---

[19] Note that others, such as Morlec, Bailly & Auberge (1997) make a diference between emotion and attitude, in that the former refers to how a speaker feels when he is uttering a sentence, whereas the latter refers to how a speaker feels about an utterance he is producing.

[20] See also Vroomen, Collier & Mozziconacci (1993), who demonstrated, through manipulation of synthesized speech, that pitch contour (along with durational variables) are sufficient to express various emotional states.

Another important aspect in the characterization of pitch is the type of discourse. Graddol (1986), for example, demonstrated that the pitch characteristics of neutral technical prose differ significantly from those of dramatic dialogue, measured in terms of mean, standard deviation, range, skew and kurtosis of fundamental frequency. According to his findings, pitch range is more extensive and pitch mean is higher in dialogue. Johns-Lewis (1986) examined pitch in three discourse modes (reading aloud, acting and conversation), performed by ten amateur actors, and found that the largest fluctuation in fundamental frequency mean is to be found in acting. Mean pitch values are also usually higher for acting, suggesting that "the heightening of pitch characteristics is attuned to attention-states in the hearer-organism, pitch heightening having the function of focusing and maintaining attention" (p. 217).[21]

Speech style also plays a decisive role on the characterization of pitch patterns. Bruce *et al.* (1997), for example, observed a greater variability in fundamental frequency maxima and more extreme values for spontaneous speech, as opposed to read and/or acted speech. Intonational prominence, or the "obtrusion of an intonational peak from median Fo across an utterance" (Cutler 1997), tends to be greater in spontaneous speech, as Koopmans-van Beinum (1990 and 1991) suggests.[22] Additionally, read material often exhibits higher mean pitch, steeper slope of declination, and stronger evidence of fundamental frequency resetting than spontaneous speech (Swerts, Strangert & Heldner 1996).

---

[21] In a study on the prosodic aspects of speech of young children, Garnica (1977) found that adults generally rise sentence-final pitch contours and add primary stresses in several places in a sentence when addressing 2-year-olds in order to get their attention.

[22] See, however, Laan (1997), who found no prosodic or spectral cues in his study that convincingly pointed to a spontaneous or read speaking style. Laan (1997: 64) proposes that "the impression of a speaking style that listeners have seems to be based on a variety of small acoustic cues".

Fundamental frequency varies as a function of gender and age of the speaker. The average fundamental frequency for women is approximately 220 Hz, for men 120 Hz, and for children 265 Hz (Cruttenden 1997).[23] Female speakers usually exhibit larger fluctuation around the mean fundamental frequency values than male speakers (Graddol 1986; Johns-Lewis 1986). According to Graddol (1986) and Graddol & Swann (1989), although the fluctuation in pitch may be explained on a physiological basis, there are social conditioning forces acting in particular ways as well. For example, according to Graddol (1986), what limits males to function within a narrower band of their pitch potential is a social factor, rather than a physiological one.

There are many examples of social factors determining fundamental frequency values in the literature. It has been reported, for example, that adults tend to use a higher pitch when they speak to young children than when they speak to adults (Phillips 1970; Garnica 1977). Pitch is also employed as a means of indicating social behaviors. In Tamil and Tzeltal, for instance, a high-pitched speech is used as a way of indicating social politeness, such as in situations where those of a low caste have to speak to those of a high caste (Brown & Levinson 1978).

According to Cruttenden (1997: 128), variation in intonation due to *style*, *class*, and *sex* is probably common in all languages. However, language universal tendencies (such as intonation-groupings and in nucleus placement) are similarly present. The similarities that can be established across languages would suggest that all aspects of intonational meaning are based on an innate foundation (see also Couper-Kuhlen 1986 for a similar argument).

---

[23] The values may range from approximately 80 Hz to 200 Hz for a typical adult male speaker, and from 180 Hz to 400 Hz for a typical female speaker ('t Hart, Collier & Cohen 1990).

There is another variable that has not been taken into account in the literature; yet, it has a decisive influence over fundamental frequency values. This variable is the acoustic phenomenon known as *intrinsic fundamental frequency* or *intrinsic pitch.* Fundamental frequency patterns of speech are often considered to be the product of an interaction between two types of variables: (i) intentional – or controlled – variables, which are language specific and act on lexical and/or utterance levels, and (ii) conditioned – or uncontrolled – variables, which are the result of co-articulation or aerodynamic interactions, and are thus beyond the active control of the speaker (Atkinson 1973; di Cristo & Hirst 1986). Intrinsic fundamental frequency of vowels, being the result of the constraint imposed by the nature of the vocal apparatus, falls into the latter category and is commonly referred to as a microprosodic effect (as opposed to prosodic, which deals with larger levels of speech).

The phenomenon of intrinsic pitch can be described as the tendency of high vowels, such as /i/ and /u/ to be characterized by slightly higher fundamental frequency values than low vowels, such as /a/ (Peterson & Barney 1952; Lehiste & Peterson 1961). Although small, this effect has been found in every language that has been so far examined for it. This would lead to the assumption that the phenomenon is universal (Whalen & Levitt 1995). The importance of accounting for the phenomenon of intrinsic pitch in any given language seems to be unquestionable. Scholars have observed that information on intrinsic pitch is indispensable for research involving the modeling of intonation contours (O'Shaughnessy 1976; Thorsen 1979; Pierrehumbert 1980; Cooper & Sorensen 1981; Shadle 1985). According to di Cristo & Hirst 1986(1986: 14), "the automatic nature of intrinsic and co-intrinsic effects could lead to a distortion of the

intended fundamental frequency pattern if compensatory processes do not intervene either during the production or during the perceptual integration of the speech." Furthermore, it has also been suggested that the segment rather than the syllable is the smallest and most basic prosodic domain (Van Heuven 1994). If this is the case, it seems to be obvious then, that intrinsic pitch effects must be taken into consideration in any prosodic analysis dealing with patterns of intonation contours. In the present research, intrinsic pitch was taken into account in board line cases where a decision of rounding up the fundamental frequency value of an intonational unit had to be made (see section 4.2.2 below for further detail on this procedure).

4.2.2. Pitch range, pitch reset and boundary tone: measuring procedures

In accordance with the methodology employed in ongoing research, three intonational features were selected for the present investigation: pitch range, pitch reset, and boundary tone. These variables are often considered to be reliable phonetic structuring devices (Swerts 1997). The characteristics and methods of measurement of each of these features will be described below.

The classification of pitch range in the literature appears to be somewhat controversial. It is often defined, in a way that is perhaps too simplistic, as being the difference between minimum and maximum fundamental frequency values (Bruce *et al.* 1997; Cosmides 1983; Van Donzel 1999). Such a definition does not take into consideration the internal variation of fundamental frequency within the range of an intonation unit. In order to account for this variation, Ladd (1996) proposed a parameter

based on two partially independent dimensions called "overall level" and "span". According to this model, pitch range is encompassed by pitch span values. Patterson & Ladd (1999) employed this parameter in their study on the relation between pitch range and the perception of speaker characteristics, and found that the best measure of span is the difference between the average non-initial accent peak value and the average post-accent valley value. Although being quite precise in capturing the range information of an utterance, Ladd's (1996) model makes use of a very time consuming and labor intensive data collection procedure, as Patterson & Ladd (1999) observe, making it practically unworkable.

An alternative measure of pitch range is the value of the fundamental frequency maximum for the intonation unit. This value is extracted from within the vowel of the syllable containing the fundamental frequency peak of the intonation unit (Menn & Boyce 1982; Nakatani & Hirschberg 1995; Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Hirschberg, Nakatani & Grosz 1995).[24] Although it may not sound plausible that the fundamental frequency value of a particular point in an intonation unit corresponds to the perception of the overall pitch of the whole unit, some researchers have already suggested that this indeed, could be the case. Cooper & Sorensen (1981) and Pierrehumbert (1980), for example, observed that the heights of the various peaks within an intonation unit tend to be the same throughout the whole unit (see also Menn & Boyce 1982 and Boyce & Menn 1979).[25] In this study, the pitch range of an intonation unit was

[24] Sometimes a more sophisticated measurement is made at the vowel's amplitude maximum. This is done because the values of the fundamental frequency peak are occasionally considered to be exaggerated, as in cases where prominence-lending pitch rises occur simultaneously with intonational boundary markers (Beckman & Ayers 1994; Swerts 1997).

[25] Some additional measures of the range has been proposed in the literature, such as those based on the standard deviation of fundamental frequency mean values (Jassem 1971; Swerts & Collier 1992) and those taken from pre-determined percentiles of the upper an lower limits of fundamental frequency values

measured as the value of the fundamental frequency maximum for that intonation unit, taken at the vowel of the syllable containing the fundamental frequency peak of the intonation unit. All the values are expressed in semitones, in order to compensate for inter-subject variability.[26]

It has been noted in several languages[27] that the fundamental frequency curve of an intonation unit tends to decline with time,[28] due primarily to physiological factors.[29] A direct consequence of this phenomenon is that the values of fundamental frequency at the end of an intonation unit have a tendency to be lower than the values at the beginning, producing an acoustic effect known as "fundamental frequency reset", or "pitch reset". Although the phenomenon described here is physiological in nature, it can partially be manipulated by the speaker for discourse reasons (Swerts & Geluykens 1994; Pijper & Sanderman 1994).[30] In the present investigation, pitch reset was calculated as the difference between the pitch range values of two adjacent intonation units.

(Graddol 1986; Lewis & Tiffin 1933). Couper-Kuhlen (1986: 92) describes pitch range in terms of pitch height and pitch width.

[26] In general, fundamental frequency (the rate at which a waveform is repeated per unit of time) is given in hertz (Hz). A hertz corresponds to one waveform cycle per second. However, as discussed in Section 4.2.1, fundamental frequency values vary as a function of several different variables, what makes inter-subject comparison cumbersome. For that reason, the semitone scale was devised (Baken 1987: 127). It was based on the fact that the perception of pitch change is logarithmic, and not linear. Listeners are more sensitive to changes at lower frequencies (Lehiste 1996). The standard musical semitone scale, which makes use of an exponential growth function, reflects this fact.

[27] Some examples are: Danish (Thorsen 1980), Dutch ('t Hart, Collier & Cohen 1990), English (Maeda 1976), French (Contini & Boé 1975), Italian (Magno-Caldognetto *et al.* 1978), Japanese (Fujisaki, Hirosi & Ohta 1979) and Russian (Svetozarova 1975). Bolinger (1978) considers it a language universal phenomenon.

[28] Although largely accepted in the literature, this phenomenon is still a controversial one (Bruce 1984; Cohen, Collier & 't Hart 1982; Gelfer *et al.* 1983; Hirst & Cristo 1998; Ladd 1984). It has been suggested that rather than being a natural process found in any circumstance, declination is much more evident in laboratory speech (Lieberman & Tseng 1981; Umeda 1982).

[29] The physiological explanations for this phenomenon vary. Some authors attribute it to the decline in transglottal pressure (Lieberman 1967), others prefer to interpret it as a result of a laziness principle (Ohala & Ewan 1973) or a trachea pull (Maeda 1976).

[30] It has also been proposed that speech style has a significant influence on declination (Cruttenden 1997; Levelt 1989; Swerts, Strangert & Heldner 1996; Umeda 1982).

The speech files were digitized at a rate of 22.05 KHz with 16-bit resolution using the speech-editing software SoundEdit 16™, version 2.0 (Macromedia Inc.). The data was subsequently analyzed under the speech-editing program Praat™, version 3.8.24 (Boersma 1997). Pitch values in the signals were extracted automatically, using the default fundamental frequency extraction algorithm in the program.[31] The original pitch contours were then stylized by hand, in a semi-automatic process that used both visual and auditive cues. This was done to avoid the interference of octave jumps and to smooth the contours (Nooteboom & Kruyt 1987; Sluijter & Terken 1993; Van Donzel 1999). Fundamental frequency peak values in the signal contours could be taken automatically from the program's information window. In the measurement, fundamental frequency values, expressed in semitones, were rounded up. The phenomenon of intrinsic pitch was taken into consideration at this time.[32] Furthermore, pitch values that were considered to be unreliable due to the presence of laryngealization or minimal intensity were disregarded from the final computation.

As described in Chapter 1, boundary tones were assigned on a perceptual basis by five experts in Brazilian Portuguese intonation. Using a method similar to that which was employed by Van Donzel (1999), verbatim transcriptions of the narratives were given to each of the five Brazilian Portuguese intonation experts, in addition to a tape containing the spoken version of all the seventeen narratives. Besides indicating the location of intonation unit boundaries, the experts were asked to specify the position of pitch accent

---

[31] The default pitch algorithm used in Praat™ 3.8.24 is based on an autocorrelation method (see Boersma 1993 for more information on the algorithm). The default values used by the program are: time step: 0.01 sec; range: 75-600 Hz; silence threshold: 0.03.

[32] In a few borderline cases where a decision of rounding up to a higher or lower value had to be made, the vowel from which the value was extracted was taken into account for the final decision. For a discussion of the intrinsic pitch phenomenon, see Section 4.2.1 above.

in each intonation unit and to label the type of boundary marking pitch movement as either "low" or "non-low".[33] In order for a boundary to be considered as "low" or "non-low" in the present work, three out of the five experts had to agree in their judgment. In general, the judgments were very consistent from experts to expert. In fact, most boundary tones were classified as either "low" or "non-low" unanimously.

4.3. Pitch in the data: an overview

Section 4.3 provides an overview of the fundamental frequency values measured in all of the narratives in the corpus, together with the fundamental frequency values from the perspective of the individuals who participated in the research.

4.3.1. General characteristics of the narratives

Table 4.1 below presents the values of fundamental frequency minimum, maximum, and mean, together with the standard deviation from the mean values. The values are given in semitones and refer to each narrative as a whole.

[33] This classification was inspired by the problems reported in Brown, Currie & Kenworthy (1980), Geluykens & Swerts (1994), Swerts, (1997)], and Van Donzel (1999) regarding the reliability in the distinction of "high" from "mid" tones. "Non-low" tones in the present study covers both "high" and "mid" tones, thus.

Table 4.1
Overview of some intonational features in the data, broken down by narrative: minimum fundamental frequency, maximum fundamental frequency, mean fundamental frequency and standard deviation. All values are expressed in semitones, computed relative to 100 Hz (0 ST equals 100 Hz), and characterize the narratives as a whole

Narrative Number	*Min. Fo*	*Max. Fo*	*Mean Fo*	*Standard Deviation*
01	-5	24	15	7
02	-4	23	8	5
03	-6	25	8	4
04	-6	25	8	5
05	-6	16	9	3
06	-5	31	16	6
07	-5	25	12	3
08	-6	23	7	4
09	-4	21	13	4
10	-3	31	18	5
11	-5	29	17	5
12	-5	30	12	5
13	-6	31	15	6
14	0	31	20	5
15	-6	24	13	4
16	-4	30	17	4
17	-4	21	11	4

The data for mean fundamental frequency values shows that the variation is significant. Narrative 14, told by a female speaker (participant 03), has a mean fundamental frequency that is not only the highest of all the narratives in the data, but is also almost three time as high as narrative 8, told by a male speaker (participant 05). This is, in a way, predictable, since, as discussed above, females generally speak in a higher pitch than males. The difference, however, is substantial, in comparison with other studies (Tielen 1992; Van Donzel 1999). The variation around the minimum fundamental frequency is very small. However, the same does not hold true for fundamental frequency maxima. The values for maximum fundamental frequency range from 16 semitones to 31 semitones, a considerably high difference of 15 semitones.

The fundamental frequency maxima will be considered as the measure of pitch range. The distribution of fundamental frequency maxima (HIFo), or pitch peak, in the data is presented below.

Table 4.2
Distribution of fundamental frequency maxima (HIFo) in the data: maximum, minimum, and mean values, accompanied by the standard deviations. All the values are expressed in semitones, computed relative to 100 Hz (0 ST equals 100 Hz), and characterize the narratives as a whole

Narrative Number	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*08*	*09*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*
HIFo Max.	26	23	25	25	17	31	25	23	21	31	29	30	31	31	24	30	21
HIFo Min.	11	6	6	1	6	11	11	3	11	12	10	12	8	16	13	14	10
HIFo Mean	20	14	14	14	11	24	16	11	16	24	22	17	22	23	17	23	15
HIFo sd	4	5	5	6	2	5	3	6	3	4	5	5	7	4	3	4	3

The values of fundamental frequency maximum across narratives range from 1 ST to as high as 31 ST. The lowest value is found in narrative 04, told by a male speaker (participant 05), while the highest value was attested to female speakers: narratives 06 and 13 (participant 08) and narratives 10 and 14 (participant 03). This renders a statistically significant effect: ($F(16,626)=43.2110$, $p<0.0001$). Post-hoc tests (Tukey-Kramer multiple comparisons) show two homogeneous subsets of narratives (narratives within one group do not significantly differ from one another): (i) 06, 10, 11, 13, 14, and 16, and (ii) 02, 03, 04, 05, 07, 08, 09, 12, 15, and 17. Narrative 01 falls between both groups, but differs significantly from narrative 10 of group (i) and all the narratives from group (ii), but narratives 07, 12, and 15. In general, as it will be made clear in the next subsection, this segmentation actually reflects fundamental frequency maxima as being a function of gender differences. All the narratives in group (i) were told by female speakers. Most of the narratives in group (ii) were told by male speakers (the only exceptions are narratives 07, 09, 12, 15 and 17). Narratives 06 and 10 are characterized by the use of the highest mean pitch peak in the data. Both stories were told by female speakers.

4.3.2. General characteristics of the participants

Table 4.3 below displays the values of fundamental frequency minimum, maximum, and mean, together with the standard deviation from the mean values. The values are all expressed in semitones and are the result of the examination of the contribution of each participant.

Table 4.3
Overview of some intonational features in the data, broken down by participants: minimum fundamental frequency, maximum fundamental frequency, mean fundamental frequency, and standard deviation. All the values are expressed in semitones, computed relative to 100 Hz (0 ST equals 100 Hz)

Participant Number	*Min. Fo*	*Max. Fo*	*Mean Fo*	*Standard Deviation*
01	-4	21	12	4
02	-5	30	17	6
03	-3	31	18	5
04	-6	30	13	5
05	-6	25	8	5
06	-6	16	9	3
07	-5	25	12	3
08	-6	31	15	7

Mean fundamental frequencies in the data, analyzed as a function of the different participants, ranges from 9 ST (participant 06, male) to 18 ST (participant 03, female). In other words, the highest mean fundamental frequency in the data, as a function of the participants, is double the value of the lowest mean fundamental frequency. This is still a considerable variation, but in this case it is comparable to van Donzel's (1999) results. Results from an analysis of variance show that mean fundamental frequency values are significantly different for the male and the female participants in the data: ($F(1,7)=9.5294$, $p>0.0215$). The variation of minimum fundamental frequency is not as large (ranging from –4 to –6, a 2 ST of maximum distance). Maximum fundamental

frequencies, however, present a large variability: 16 ST (speaker 06, male) to 31 (speakers 03 and 08, females), a 15 ST difference. The distribution of fundamental frequency maxima as a function of the participants is given in Table 4.4 below.

Table 4.4
Distribution of fundamental frequency maxima (HIFo) in the data, broken by participants: maximum, minimum and mean values, accompanied by the standard deviations. All the values are expressed in semitones, computed relative to 100 Hz (0 ST equals 100 Hz)

Narrative Number	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*08*
HIFo Max.	21	30	31	30	25	16	25	31
HIFo Min.	10	10	12	12	1	6	11	8
HIFo Mean	16	23	24	17	13	11	16	23
HIFo sd	3	5	4	4	6	2	3	6

With the exception of the two male speakers in this research (participants 05 and 06), particularly participant 05, the lowest values of pitch peak for most speakers center around 10 ST. As for the maximum values, a considerable difference of 10 ST among the female group can be verified (participants 01 and 03/08). The lowest peak value in this category is that of participant 06, a male speaker (17 ST). Participants 05 (male) and 07 (female) follow, making use of exactly the same value of maximum pitch peak values: 25 ST. As for the mean values of maximum fundamental frequency, a difference of 13 ST is verified between speakers 06 and 03, a considerably large value. Analysis of variance shows that participants differ from each other in terms of mean pitch peak values: ($F(7,626)=92.5925$, $p<0.0001$). Post-hoc tests (Tukey-Kramer multiple comparisons) show three homogeneous groups of participants (participants within one group do not differ significantly from one another): (i) 02, 03, and 08; (ii) 01, 04, and 07; and (iii) 05 and 06. Groups (i) and (ii) are formed by female speakers; group (iii) is formed by male speakers only. This result clearly demonstrates that gender plays an important role in the

characterization of discourse in terms of fundamental frequency values. Further statistical analysis substantiates this assumption. Mean fundamental frequency values for the male and the female speakers differ significantly ($t=17.441$, $df=625$, $p<0.0001$).

4.4. Pitch cycle

One of the main objectives of the present dissertation is to identify the role of prosodic features in the segmentation of narrative discourse. It was demonstrated in the previous chapters that both pause and speech rate play important roles in the deliniation of narrative structure. A linear analysis of these variables indicated the existence of a cycle of varying values that corresponded to the way narrative texts are generally organized. Since the investigation of this phenomenon in the narratives as a function of different temporal phenomena showed a positive correlation, the examination of the same phenomenon in terms of fundamental frequency value seems to be justifiable.

An examination of the mean values taken from each individual section of all the narratives in the data revealed that the pattern found for the temporal variables could not be obtained in terms of fundamental frequency values as well. Mean pitch values do not follow any specific pattern in the narratives under analysis, if considered from a linear perspective. The following examples illustrate this "haphazard" aspect of the linear pitch distribution in the narratives:

Figure 4.1
Fundamental frequency values plotted over time for narrative 03

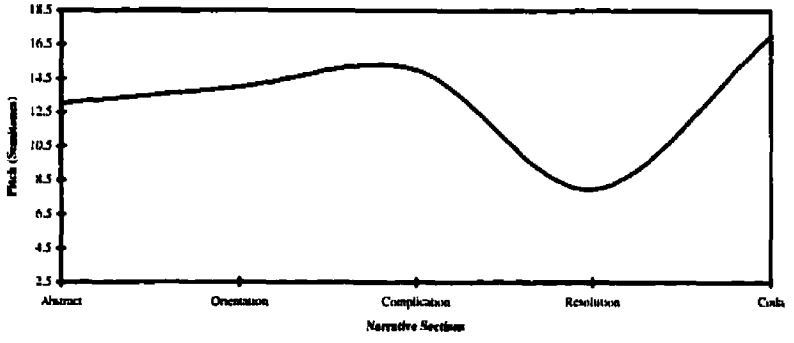
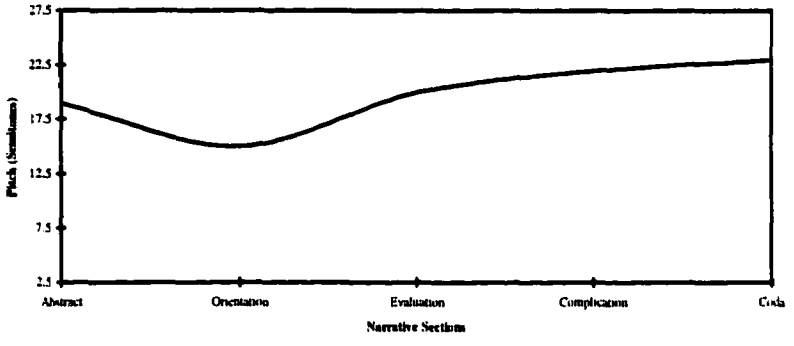


Figure 4.2
Fundamental frequency values plotted over time for narrative 01



The main pitch range value in the abstract section of narrative 03 (Figure 4.1) is 13 ST. It increases slightly in the orientation section and goes up to 15 ST in the complication section. The pitch range then drops to almost half of its value in the resolution section and finally reaches its highest mean value, 17 ST, in the coda of the narrative. No alternation of high and low values characterizes this cycle, and apparently, no consistent pattern can be established here, as observed in connection with pausal phenomena and speech rate. A comparison of the pattern found in narrative 03 with the one displayed in Figure 4.2 for narrative 01 clearly suggests that the distribution of pitch range values in narratives, from a linear perspective, is not determined by structural

factors. In narrative 01, mean pitch range starts at 19 ST and, instead of increasing, as it did in narrative 03, it drops 3 ST in the orientation section. It increases again in the evaluation section to almost the same value of the abstract, suggesting the existence of a possible "cycle" here; however, instead of decreasing again, the pitch range value in the proceeding section rises 2 ST and continues rising until the last section of the story, resulting in a completely different pattern from the one derived from narrative 03 (Figure 4.1). This inconsistency characterizes all the remaining narratives in the data. Apparently, speakers do not seem to exploit pitch range in the same way they do with speech rate and pausal phenomena lay out the structure of their stories.[34] The way pitch range is employed in narratives is most likely the result of other factors, one of which will be examined next.

In the following section, pitch range distribution will be considered as a function of individual narrative sections in order to verify whether a correlation between pitch range values and semantically individualized story parts (narrative sections) can be established. It will be argued that narrative structure, rather than being outlined in terms of prosody by a linear variation in pitch range values, is actually attested through the characterization of narrative section in terms of pitch range values.

4.5. Pitch range in narrative sections

It is often assumed in the literature that pitch is used to highlight the parts that are considered to be the most important in a particular discourse (Cutler & Ladd 1983: 142;

[34] Douglas-Cowie & Cowie (1998) found that particular sections in a corpus of telephone conversations of secretaries could be distinguished on the basis of Fo values, among other things.

Bolinger 1978: 515). This phenomenon can be explained from a number of different perspectives, the most common being the so-called given/new distinction, as discussed above (see, among others, Chafe 1976; Clark & Haviland 1977; Eady *et al.* 1986; Halliday 1967), or the foreground/background distinction, from a prosodic standpoint (Kumpf 1987; Reinhart 1984). In general terms, what is considered to be significant, new or foregrounded information in speech is often put in focus, and this is acoustically achieved by means of a pitch accent. Given, irrelevant, or backgrounded information, on the other hand, tends to be not focused, being thus often uttered in an average or low pitch.

Most of these studies, however, restrict their analysis to a local level, i.e., they focus on the information status of the individual words within an intonation unit, rather than on the relation that these intonation units have among each other, in a more global level.[35] On the local level, focus is for the most part prosodically signaled by means of a pitch accent (Baart 1987; Currie 1980; Eady & Cooper 1986; Nooteboom & Kruyt 1987; 't Hart & Collier 1975; Terken 1984).[36] According to Brown, Currie & Kenworthy (1980: 76), pitch range is also consistently used "to indicate the relative importance of items in terms of information value". By comparing the values of fundamental frequency measured in recurring lexical items and in unstressed syllables/stressed grammatical items, Brown and her associates found that pitch range is employed not only to indicate whether a word contains given or new information in the discourse, but also to

[35] Note that sometimes a group of words can be put in focus. This is done by means of an accent on a single word, which is often called the "prosodic head" of the constituent (Nooteboom & Kruyt 1987). Again, the assignment of focus is restricted to a local level, i.e., within the intonation unit domain.

[36] The acoustic correlate of pitch accent is sometimes considered to be determined by the extent of fundamental frequency changes, and other times to be a function of fundamental frequency maximum. Terken (1991) agues that neither explanations account satisfactorily for the actual perception of prominence of accented syllables, and proposes that local and global characteristics of intonation phrases must be taken into account for an accurate phonological rule designed to predict prosodic prominence.

discriminate between the amount of information carried by different kinds of non-lexical items.

Couper-Kuhlen (1986: 198) proposes that, on the global level, focus can be identified by means of pitch height and/or pitch range. According to her study, the focus of a minor paratone will most likely fall on the word with the greatest pitch height and/or pitch range of all the other words occupying the nuclear position in the same paratone.[37] Couper-Kuhlen (1986), however, considers the pitch range value taken at a given lexical element as the measure of pitch range for the entire discourse unit.

No attempt has been made so far, to the best of my knowledge, to correlate pitch range to overall discourse segments (or paragraphs, as some authors prefer[38]), by taking into account the message that such units convey. Based on the assumption that the pitch height/range is often associated with general textual meanings, the present section will investigate whether a significant correspondence between mean pitch range on the global level can be verified. The hypothesis to be tested is that sections of a narrative that are considered to be essential will present a mean pitch range higher than the average pitch range of the entire narrative. The corollary of this hypothesis is that narrative discourse does present a well-defined structure, and that this structure is evinced in speech by means of prosody.

In general terms, only the complicating action and the evaluation sections are fundamental in a narrative (Labov 1972; Labov & Waletzky 1967). However, as it was observed earlier, occasionally some other sections in a narrative present a large number of linguistic features regularly used as evaluative devices. This most often results in that

[37] For a discussion of the notion of "paratones", see Yule (1980).
[38] For example, Hinds 1979, Sluijter & Terken 1993, and Nakajima & Allen 1992.

section incorporating prosodic properties characteristic of evaluations. It was suggested in Section 2.10, Chapter 2, that this is often the case of coda sections. Codas frequently sum up the point of a narrative and contain general comments accompanied by all sort of evaluative features. For that matter, it was hypothesized in Chapter 2 that codas, along with evaluations, would form a single category of sections characterized by the more complex process of interpreting.

In following the Labovian criteria of narrative sections importance, only complications and evaluations are characterized as essential. In that sense, it is hypothesized here that complication and evaluation sections will present a larger pitch range than abstracts, orientations and codas. Resolutions, though not being essential in a narrative, will be grouped with complications and codas, since they share the same formal properties with complicating sections.

Before the general results are presented, the pitch range behavior will be analyzed in detail within a single narrative text. The analysis reveals certain aspects of pitch range at the local level in ways that a more global analysis cannot account for. For this purpose, narrative 14 was chosen. This narrative tells how the speaker got in trouble when she was in a ski station, riding the ski lift. The story was narrated by the same participant who told narrative 10 and is actually a continuation of the topic initiated in narrative 10: amusing accidents in the ski station. Table 4.2, which brings the results of the measurements of pitch range in all the narratives, reveals that the mean value of pitch range for narrative 14 is 23 ST. The average fundamental frequency for this narrative was 20 ST, the largest in the data. For the purposes of the analysis below, any intonation unit displaying a range above this mean value is considered to be marked, having thus a special status in relation

to the others. These 'marked' values will appear in bold face in the transcription below. The purpose of this analysis is to provide an analysis of pitch range behavior in an entire narrative, rather than focusing on trends common to the whole data.

Abstract	
01 – ai outra experiência	**24**

Abstract	
01 - *oh another experience*	**24**

The abstract of this narrative is restricted to one intonation unit only. This is characterized by a pitch range larger than the average. As mentioned earlier, this narrative is a second illustration of accidents at the ski station, a topic that had been previously discussed. The expanded pitch range, together with a (perceptually) higher intensity, and accompanied by the interjection at the beginning of the utterance, reflects the sudden recollection of another *experience* similar to the one described in narrative 10. It may also be interpreted as a strategy that is employed to indicate that a story is about to begin. In this sense, a larger pitch range at the beginning of a narrative is expected, because it signals "the story-narrator's control of subsequent verbal behavior" (Brazil 1997: 93).

Evaluation	
02 - bom	18
03 - aí depois de um tempo a gente aprendeu	**28**
04 - eu comecei a aprender né?	**27**
05 - fiquei com mais facilidade	19
06 - e eu sempre ia na verde	21
07 - que é mais fácil e tal	16
08 - me divertia	21

Evaluation	
02 – *well*	18
03 - *then after a while we learnt it*	**28**
04 - *I started to learn it*	**27**
05 - *I found it easier*	19
06 - *and I always went in the green lane*	21
07 - *which is the easiest and all*	16
08 - *I had a good time*	21

This section is primarily evaluative, because it establishes the point of the narrative. This could be paraphrased as: "look how stupid and funny I am: even after learning how to ski, I made a foolish mistake for the second time that called everybody's

attention to me".[39] However, this section also contains a great deal of information that serves to orientate the audience. For example, it gives information about the time (*after a while*, i.e., after the first accident, described in narrative 10) and the location (*the green lane*: this narrative is still situated at the ski area, something that is only suggested in the abstract). As discussed earlier, orientation sections are not generally presented in expanded pitch range, simply because they convey information that is not as essential as the information carried on in the complication and evaluative sections. The mean pitch range of this evaluative section is 21, which is lower than the average range for the whole narrative. However, two intonation units in this section are characterized by a noticeably expanded range (IUs 03 and 04). Note that these two utterances convey central evaluative information that is directly linked to the point being made. What the narrator implies here is that even after learning how to use the ski station facilities (*we learnt it...I started to learn it*), she made another foolish mistake, contributing to the image of herself she wants to portray: a clumsy person.

Complication	
09 - até que um dia/	**30**

Complication	
09 - *when it all happened*	**30**

The beginning of the complication section is intonationally marked by the use of an extremely wide pitch range. Note that this section is interrupted by an orientation section conveying information intended to facilitate the comprehension of the described events. The brief delay of the narrative stream, rather than considered to be a negative aspect of the telling, actually increases the suspense of the story and for that matter can be regarded as an evaluative device.

[39] The particularities of narratives that supposedly put the storytellers in a negative position is discussed in Oliveira (1999).

Orientation	
10 - bom você desce né?	23
11 - e quando você sobe/	18
12 - é assim	18
13 - quando você sobe éh/	20
14 - como se fosse uma coisa assim	22
15 - então você vai/	17
16 - você chega na plataforma	19
17 - os teus esquis encostam no chão	17
18 - e daí tem uma rampa	22
19 - e você vai pra rampa	23
20 - e as cadeirinha continua pra cima	21
21 - e você tem que descer aqui	23

Orientation	
10 - *well you go down right?*	23
11 - *and when you go up/*	18
12 - *it's like that*	18
13 - *when you go up er/*	20
14 - *it's like it was something like this*	22
15 - *then you go/*	17
16 - *you reach the platform*	19
17 - *your skis touch the floor*	17
18 - *and then there is a ramp*	22
19 - *and you go to the ramp*	23
20 - *and the chairs continue their way up*	21
21 - *and you have to get down here*	23

This orientation section contains no utterance displaying a pitch range larger than the average. It simply brings information on how the ski lift system works. The message is fragmented and the whole section is accompanied by a number of extra-linguistic features, such as hand movement, helping to illustrate the 'patchy' atmosphere that the section bears. Although the information conveyed in this section could be regarded as important, because it clarifies the operation of the ski lift, it is not essential to the narrative. For that reason, it is told in an average pitch range.

Complication	
22 - e eu não sei o que aconteceu que eu não consegui descer	**28**
23 - em vez de eu descer na rampa	**27**
24 - eu conti/ fiquei na cadeira	**28**
25 - claro que a cadeira parou de novo né?	22

Complication	
22 - *and I don't know what happened I couldn't get down*	**28**
23 - *rather than going down to the ramp*	**27**
24 - *I conti/ I stuck in the chair*	**28**
25 - *of course the chair stopped again*	22

After the long and fragmented orientation section, the complicating action of the narrative restarts. Once again, it is intonationally marked by the use of a pitch range larger than the average.[40] The mean pitch range of the entire section is 26 ST. Most

[40] It should be noted here that the intonation units displaying values higher than the average are considered to be marked from a conceptual perspective only. It is expected that these values are above a perceptual threshold, being thus *perceptually* marked too. However, since the present dissertation does not test for perception, such a claim cannot be made.

intonation units in it are produced in an expanded pitch range. The only exception is IU 25, which does not contribute any new information, rather, it is solely an obvious conclusion of the facts.

Evaluation	
26 - um sábado	**24**
27 - todo mundo olhando pra mim	**28**

Evaluation	
26 - *a Saturday*	**24**
27 - *everybody looking at me*	**28**

This evaluative section, as opposed to the previous one, is characterized by a mean pitch range larger than the average: 26 ST. It is noteworthy that this evaluation is the one that most explicitly makes the point of the narrative: the speaker once more became the center of attention because of her nonsensical behavior. Both intonation units that compose this section are told in a range larger than the average.

Resolution	
28 - vieram eu acho que umas cin/ umas cinco pessoas	**29**
29 - tiveram que fazer pezinho assim pra mim descer	**30**

Resolution	
28 - *about fi/ I think about five people came*	**29**
29 - *they had to help me with their hands like this*	**30**

The resolution section is in synchrony with the complication section in terms of pitch range. Resolutions are not essential in narratives, as Labov (1972) pointed out. In many cases, a narrative can come to an end without having a formal resolution section. However, if the resolution is present in a story, it usually carries important information. The events described in the resolution section of this narrative are not only important to the comprehension of the story itself, but also to the point that the narrator wants to make. Observe the extremely large pitch range in IU 29, when the speaker says that the five

people who came to help her had to "fazer pezinho"[41], an expression in Portuguese that describes a somewhat humiliating action, especially for an adult.

Evaluation	
30 - as pessoas não acreditavam	**30**
31 - fiquei com tanta vergonha	19
32 - quase morri	21
33 - mas é/... foi divertido	**31**

Evaluation	
30 - *the people didn't believe*	**30**
31 - *I got so ashamed*	19
31 - *I almost died*	21
32 - *but it is/... was funny*	**31**

The last evaluation section of this narrative presents an interesting use of pitch range. IU 30 is clearly evaluative, because it describes an impression that the narrator had about the people who witnessed her actions. The expanded pitch range in this utterance is probably an attempt to convey their astonishment. The next two intonation units make reference to the narrator's attitude towards her own actions. An abrupt drop of pitch range accompanies both utterances. This, together with a perceptual drop of intensity, may be regarded as a reflection of the shame the narrator confess she felt.[42] Finally, the last IU in the section goes back to the third person and makes a general statement about what happened. It is important to note that the mean pitch range of the entire section is 25 ST, i.e., above average, thus highlighting the section as a whole. The distribution of pitch range within the section, however, seems to be governed by locally determined semantic factors. In this specific case, there appears to be a split between the attitude of the narrator towards the events she reported and the attitude of the other characters in the narrative as interpreted by the narrator herself. All four utterances are evaluative in the sense that they contribute to the point being made; however, the way pitch range is employed here suggests that not only expanded pitch range values have an evaluative

[41] This expression is equivalent to the English expression "to cradle".
[42] Cf. Fónagy's (1978) study on the relationship between pitch and various emotional states. According to him, negative feelings are often expressed in a low pitch.

function. Low values may also be employed for this purpose, as in the present case, and only context determines their usage. Still, as mentioned above, there seems to be an attempt to intonationally single out the section as a whole. Since low pitch range was employed in the section for evaluative reasons, as proposed earlier, intonation units that were supposed to be uttered in a larger pitch range received very high values, suggesting the existence of a possible compensatory effect.

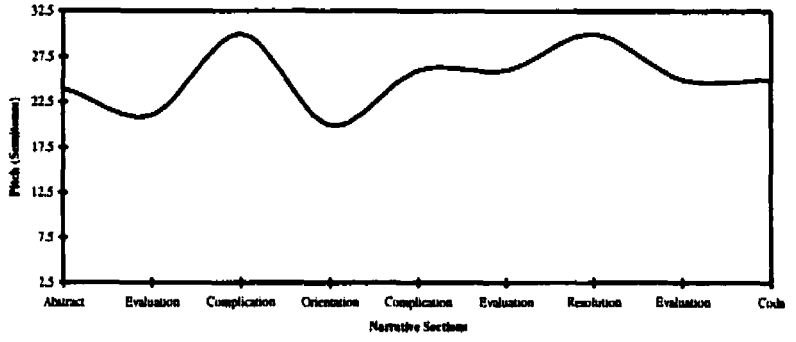
Coda	
34 - mas depois de um tempo a gente acaba/ né?	**30**
35 - agora eu já me acostumei com todos os problemas	**26**
36 – já tô indo bem	**24**
37 - então é fun	20

Coda	
34 - *but after a while we become/ you know?*	**30**
35 - *now I'm already used to all the problems*	**26**
36 - *I'm doing all right*	**24**
37 – *so it's fun*	20

The coda of this narrative, equally large in mean pitch range (25 ST), returns to the present moment. The high value assigned to this coda can be explained by taking into consideration its evaluative content. As previously mentioned, coda sections are sometimes characterized by the general assessment they regularly make of the events reported in a story, occasionally making them closely related to evaluative sections (see discussion of Chapter 02, Section 2.10). Furthermore, it is important to mention that this section starts high in terms of pitch range, but gradually decreases, until the lower-than-average value of 20 ST. This fading-away could be interpreted as another strategy, along with the use of long pauses (see discussion of Chapter 2, Section 2.10), employed by the speaker to indicate that the narrative about to end.

The distribution of pitch range in narrative 14, from a global perspective, is represented in the following figure:

Figure 4.3
Fundamental frequency values plotted over time for narrative 14

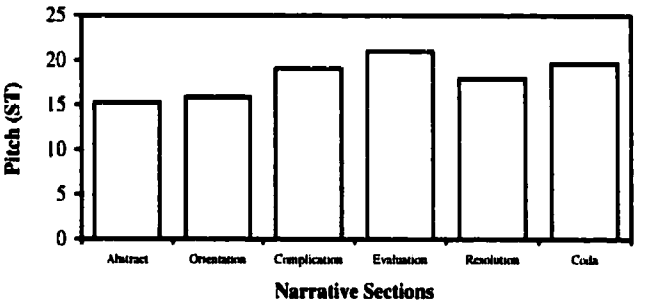


Compared to narratives 01 and 03, plotted in Figures 4.2 and 4.1 respectively, narrative 14 follows roughly the same principle that was established for those narratives: no cycle of high and low values can be found here,[43] but rather a pattern which is actually the result of semantically determined factors. In all three figures, the complication and the evaluation sections display higher values. Orientations and abstracts are always compressed in pitch range. Codas grouped with complicating actions and evaluations, displaying an expanded range. The only exception in these three cases is the resolution section of narrative 03 (Figure 4.1). Narrative 14 on the other hand has a resolution section with the most compressed value of all.

All that remains now is to confirm whether these few examples reflect a general trend in the corpus. Figure 4.4 pictures the mean values of pitch range in all the narratives of the corpus.[44] All values are expressed in semitones, computed relative to 100 Hz (0 ST equals 100 Hz).

[43] The actual values are: abstract (24 ST), evaluation (21 ST), complication (30 ST), orientation (20 ST), complication (26 ST), evaluation (26 ST), resolution (30 ST), evaluation (25 ST), and coda (25 ST).
[44] Standard deviation values were, for each section: 6.3 ST (abstract), 5.5 ST (orientation), 6.7 ST (complication), 6.1 ST (evaluation), 7.3 ST (resolution), and 7.2 ST (coda).

Figure 4.4
Distribution of pitch range in each of the six narrative sections in the corpus



As expected, both the evaluation and the complication sections have a mean pitch range larger than most of the other sections. The abstract and the orientation sections are characterized by more compressed ranges. Resolutions and codas seem to belong to the group of complication and evaluation in different degrees. Narrative sections are significantly different from each other, according to results derived from analyses of variance: ($F(5,626)=11.9539$, $p<0.0001$). Post-hoc tests (Student's t-tests) revealed that narrative sections can be separated into three different groups, according to the size of the pitch range they present: (i) large pitch range values: evaluations and codas; (ii) intermediary pitch range values: complications and resolutions; and (iii) reduced pitch range values: orientations and abstracts.[45] The fact that evaluation presented the largest pitch range in the data corroborates its primacy over the other sections, something Labov (1972) and Labov & Waletzky (1967) have suggested. As for coda sections, the above results indicate that they are in fact very much related to evaluations.

Pitch range, then, is demonstrated to be a reliable prosodic feature in terms of differentiating semantically independent sections in a narrative. This clearly contributes

[45] Codas do not differ significantly from complications and resolutions; resolutions do not differ from orientations and codas.

to the segmentation of narratives. What remains to be investigated is the extent to which pitch phenomena have a role in the segmentation of narratives. Some authors have suggested that pitch range is closely associated to discourse structure on the local level (Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Brown, Currie & Kenworthy 1980; Silverman 1987; Avesani & Vayra 1988; Ayers 1992).

In their analysis of intonational features in discourse structure, Hirschberg & Grosz (1992) and Grosz & Hirchberg (1992) investigated the role of pitch range, among several other prosodic features, in individual intonation units according to their position on the global level. Each intonation unit was classified as segment initial (*SBEG*), if located in the beginning of a topic unit, segment final (*SF*), if located at the end of a topic unit, or segment medial pop (*SMP*), if located in a medial position that did not correspond to either initial or medial position in the discourse. By comparing these three different groups, they found that intonation units located at medial position (*SMP*) and those initiating a topic unit (*SBEG*) are often uttered in a more expanded pitch range than intonation units located at the end of a discourse unit (*SF*).[46]

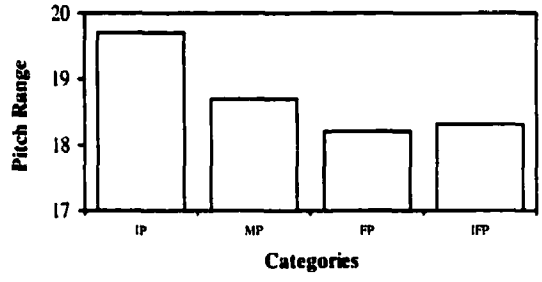
In order to verify whether pitch range can help in the segmentation of narrative texts into semantically identifiable narrative sections, the following procedure was carried out. First, intonation units were classified according to their position and function in a narrative text. Second, four categories were proposed: initial position (IP), which corresponded to those intonation units that were located at the beginning of a narrative

[46] Observe, however, that these results were originated from measurements taken from the labeling of two different groups: one that had only access to the transcriptions of the narratives (Group T) and another that had access to both the transcriptions and the speech material (Group S). In the case of *SBEG* utterances, a disparity in terms of the different groups could be verified for the measurement of pitch range. Pitch range had no effect on Group S judgments of *SBEG*. The consideration of a broader category identifying 'segmentation shifts', which includes both *SBEG* and *SMP* utterances (*SBEG* + *SMP*), yielded similar results for groups S and T.

section; final position (FP), which corresponded to the last intonation unit in a narrative section; middle position (MP), those intonation units that did not correspond to either initial position or final position. Finally, a special category was created to accommodate cases in which a single intonation unit corresponded to a narrative section. In such cases, the sentence functioned not only as the first utterance of the section, but as the last one as well as. The fourth category was then named initial-final position (IFP).

Figure 4.5 below shows the mean values of pitch range for each of the above-mentioned categories:[47]

Figure 4.5
Mean pitch range values for different intonation units classified according to its position in the narratives



A trend similar to that of Hirschberg & Grosz's (1992) and Grosz & Hirschberg's (1992) results is verified in the present study. Intonation units in initial and medial position are generally characterized by a larger pitch range than those in final position. The difference, however, is very small (especially in the case of MP utterances): statistical analyses yielded no significant effect.

The present section demonstrated that pitch range is significantly correlated to semantically individualized story parts (narrative sections). The mean range of different

[47] Standard deviations are, for each category: 6.7 ST (IP), 6.5 ST (MP), 6.9 ST (FP) and 7.3 ST (IFP).

narrative sections differs from one another according to the degree of importance the section has to the story. Apparently, speakers signal which parts in the narrative the hearer should pay more attention to by means of pitch range. This device has a secondary function, which involves making the narrative structure more transparent in terms of its own internal organization. Another possible function of pitch range was also investigated here. Based on the assumption that utterances located at the beginning of larger discourse units are often characterized by a more expanded pitch range, a comparison of intonation units in terms of pitch range was carried out in order to find out whether the position they occupy in a narrative, as analyzed according to the Labovian model, determines the size of their pitch range. Although a trend in the hypothesized direction could be easily verified, statistical analyses revealed that the differences in range as a function of the position of the intonation unit in the story are not significant.

This is not to say however, that pitch does not contribute to narrative segmentation on the local level. What will be suggested in the following section is that the *difference* in pitch range between two adjacent intonation units, rather than the pitch range value of the intonation unit, is the most important clue for the segmentation of narratives.

4.6. Pitch reset as a cue for narrative segmentation

In a study on the function of speech melody in the characterization of the sense of "finality" in utterances, Swerts, Bouwhuis & Collier (1994), use a perceptual approach to compare different melodic variables. They found that variations in pitch register, pitch

contour, and pitch range are important cues in the signaling of the rounding off of a particular discourse unit. Although all three parameters demonstrated to contribute significantly to the listener's judgment of finality, a disparity in the strength of these cues could be observed: the differences in register and contour were equally effective for the perception of finality in utterances, but the impact of pitch range was considerably lower than the other two features.

It has been suggested that the melodic discontinuity that occurs between information units - a consequence of the natural declination of pitch in the course of an utterance[48] - is an important cue for discourse segmentation (Pijper & Sanderman 1994; Swerts 1997; Nakajima & Allen 1992; Hakoda & Sato 1980; Hirschberg & Grosz 1992; Grosz & Hirchberg 1992). Grosz & Hirchberg (1992), for example, observed that automatically generated prediction trees (produced by means of labelers' consensus decisions on discourse structure), distinguish intonation units that occur at the end of a discourse segment from other units in 92.5% of cases studied. They were employing a simple combination of subsequent pause, overall speech rate and amount of fundamental frequency change from prior intonation units. Swerts (1997), on the other hand, found a low correlation between the boundary strength values derived from labelers' consensus decisions and the distance between the pitch range values before and after a given boundary of a particular strength.[49] This measure accounted for only 17% of the variance obtained from a linear model fitting procedure that used a method of least squares regression.

---

[48] See discussion of Section 4.2.2.

[49] For a discussion of the notion of *perceptual boundary strength* (PBS), see Pijper & Sanderman (1994).

The present section will investigate to what extent the "pitch reset" (i.e., the difference in pitch range between two adjacent intonation units) contributes to the segmentation of narrative texts. It is hypothesized that in storytelling, intonation unit boundaries that coincide with a narrative boundary will present a higher pitch reset value than those that do not coincide.

Before presenting the statistical results, some examples will be examined in detail. Figure 4.6 below displays the values of pitch reset between all intonation units in narrative 16. The values are displayed linearly, and broken down by narrative sections.

Figure 4.6
Pitch reset between intonation units of narrative 16. All values are expressed in semitones

1. Complication	4							
2. Evaluation	4							
3. Complication	0	2						
4. Evaluation	1	0	6	2	13			
5. Complication	4	1	4	6				
6. Evaluation	3	4	7	9				
7. Complication	2	7	6	2	2	4	1	12
8. Evaluation	8	4	3					
9. Coda	4	5	*					

Clearly, a correspondence between the position of the intonation unit and the value of pitch reset can be established in the above example. The last intonation units in most of the sections of narrative 16 that have more than one intonation unit are characterized by the highest pitch reset values. The only exception occurs in the last evaluative section in the narrative, for a reason that will be considered in detail in a later section.[50] It is noteworthy that high values are not restricted to boundary positions only. Some transitions that do not coincide to narrative boundaries do present high pitch reset

[50] The reset value that occurs at the end of the narrative is not being considered, since it interacts with an intonation unit that is not part of the narrative itself.

values. What needs to be pointed out is that, in most cases, high pitch reset values coincide with narrative boundaries. Moreover, it should be also observed that the two highest pitch reset values in this narrative occur exactly at the end of a narrative section, i.e., at a boundary site.

The last evaluative section in narrative 16, as already stated, is the only exception in terms of the linear distribution of pitch reset. The highest value that this section presents is located right after the first intonation unit in the section, rather than after the last one, as expected. The reason for this lies in the way pitch range is manipulated by the speaker. The following excerpt brings the last two intonation units of the last complication section of narrative 16, as well as the whole evaluation section that follows it. "Rg" stands for "Pitch Range" and "Rt", for "Pitch Reset". All values are expressed in semitones.

Narrative 06	Rg	Rt
Complication		
24 - ele me liga	30	1
25 - dizen/ uma desculpa amarela	29	12
Evaluation		
26 - porque ele queria/	17	8
27 - lógico que arrumou outra perua	25	4
28 - ir com né? com ela na festa	21	3

Narrative 06	Rg	Rt
Complication		
24 - *he calls me*	30	1
25 - *saying/ a lame excuse*	29	12
Evaluation		
26 - *because he wanted/*	17	8
27 - *obviously he found some other bitch*	25	4
28 - *to go with right? with her to the party*	21	3

The mean pitch range of the last complication section in narrative 16 is 25 ST. The mean pitch range of the following evaluation section is 21 ST. The average for the entire narrative is 23 ST (see table 4.2 for further details). The pitch range that characterizes IU 26 is well below all of these values, and diverges from the surrounding values as well. This could be interpreted as a deliberate attempt to mark a narrative section intonationally, through the use of a large pitch range reset.

As previously stated, pitch range also varies as a function of the linking of utterances. According to Hakoda & Sato (1980), whenever two neighboring phrases are closely inter-connected, speakers tend to suppress the pitch peak of the second phrase more relative to the first, in order to indicate their proximity. It follows then, that the opposite may also be true, i.e., whenever a speaker wishes to indicate that two utterances are *not* interconnected, s/he may manipulate the pitch range of the second utterance relative to the first, by suppressing or even expanding its size, thereby making the difference between them larger.[51]

Two additional examples that illustrate this point are provided below. The first one was extracted from narrative 08:

Narrative 08	Rg	Rt
Abstract		
01 - um dia meu irmão chegou de cueca em casa só	5	5
Orientation		
02 - tava passando no parque ali	10	5
03 - foi dois mo/ dois moleques lá	5	2
04 - roubaram tudo dele	7	0
05 - tava voltando do inglês	7	1
06 - chegou em casa chorando	6	8
Orientation		
07 - daí tava eu e meu outro irmão em casa	14	10

Narrative 08	Rg	Rt
Abstract		
01 - *my brother came back home wearing only his pants*	5	5
Orientation		
02 - *he was passing by the park over there*	10	5
03 - *two bo/ two boys there*	5	2
04 - *they stole everything from him*	7	0
05 - *he was coming back from his English class*	7	1
06 - *he arrived at home crying*	6	8
Orientation		
07 - *so there we were me and my other brother at home*	14	10

As discussed in Section 4.5 above, both abstracts and orientation sections are, in general, characterized by a more compressed pitch range. The mean pitch range of narrative 08 is 11 ST. The orientation section, which is subdivided into two individual portions, presents a mean range of 8 ST. The narrative starts in a very narrow pitch range

[51] The value is actually expanded or suppressed in relation to the average pitch range employed in the whole discourse, which would obviously reflect the pitch range of the speaker. The value of the first utterance serves as a reference point.

(5 ST), but changes abruptly to a wider range (10 ST) when there is a change to another section. The value drops again to its previous value (5 ST) and varies just slightly (around 1 ST) until the advent of a new section (or a new portion of the same section). Once again, the pitch range becomes much wider (14 ST), diverging considerably from the average rate of the section to which it belongs. In contrast to the previous example, the larger pitch resets in this example are the result of an expansion of the pitch range in the second utterances of the pairs, rather than a compression of them. In both cases, however, a clear consequence of the manipulation of pitch range at the local level can be easily observed. The pitch reset values in other neighboring boundaries that do not coincide with a narrative boundary are also altered. This is the case of IU 26 of narrative 06 and IU 02 of narrative 08. Both utterances initiate a narrative section and, although they do not concur with the ends of sections, they are nonetheless accompanied by a high pitch reset value, due to the manipulation of their pitch range.

The last example to be analyzed here is an excerpt from narrative 11. It presents the first evaluative section of the narrative, together with the complication action:

Narrative 11	**Rg**	**Rt**
Evaluation		
15 - não sei que é que a ge/ que a gente foi fazer lá	24	1
16 - porque eu só ia em lugar porra louca entendeu?	25	2
17 - eu só freqüentava (0.34) o o lugar que Arrigo Barnabé freqüentava em São Paulo tal sabe?	23	1
18 - esses lugar	22	6
19 - e a gente foi nesse bar que é/ né?	28	9
20 - nessa boate sei lá que/	19	5
21 - não sei da onde que veio a idéia	24	1
22 - que nós fomos lá	24	7
23 - enfim	17	7
24 - uma turma	10	14

Narrative 11	**Rg**	**Rt**
Evaluation		
15 - *I don't know what we/ what we went to do there*	24	1
16 - *because I usually went to crazy places you know?*	25	2
17 - *I usually went to* (0.34) *those those places like the ones Arrigo Barnabé went in São Paulo you know?*	23	1
18 - *places like those*	22	6
19 - *and we went to that bar that/ you know?*	28	9
20 - *that club I don't know that/*	19	5
21 - *I don't know where this idea came from*	24	1
22 - *that made us go there*	24	7
23 – *anyway*	17	7
24 - *a bunch of people*	10	14

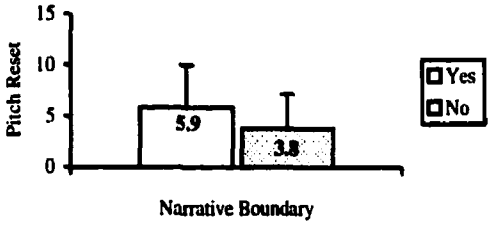
Complication		
25 - aí eu tô tomando/	24	4
26 - acho que era um uísque que eu tava tomando assim	28	6
27 - eu ponho/	22	4
28 - eu vou por o copo na mesa	26	1
29 - vem um cara assim por trás	25	3
30 - me segura aqui pela cin/ cintura	22	2
31 - eu não sei	20	6
32 - toma meu co/	26	3
33 - não eu sei que ele toma meu copo assim	29	1
34 - me vira e sai dançando	29	4

Complication		
25 - *then I'm drinking/*	24	4
26 - *I think it was a brandy that I was drinking like this*	28	6
27 - *I put/*	22	4
28 - *when I was going to lay the glass on the table*	26	1
29 - *some guy came like this behind me*	25	3
30 - *he holds me by the*	22	2
31 - *I don't know*	20	6
32 – *he took my gla/*	26	3
33 - *I don't know how he took my glass like this*	29	1
34 - *he turned me over to him and started to dance*	29	4

A quick review of the excerpt above reveals that the same strategy that is employed in the previous examples is at work here as well. The average pitch range of the first evaluation section in this narrative is 22 ST. The complication section has a 25 ST average pitch range. The narrowest pitch range in this excerpt occurs at a narrative boundary site, i.e., at the last IU of the evaluation section (IU 24). In the previous examples, it was demonstrated that the manipulation of pitch range for segmentation reasons tends to occur in the *first* IU of the following narrative section. This example illustrates that it is also possible to manipulate the pitch range of the *last* IU of a section as well, in order to widen the range between the IU and the subsequent one.

The above analyses indicate that there is indeed a strong relationship between pitch range reset and narrative boundary. Figure 4.7 below presents the average values of pitch reset, measured at different boundary types.

Figure 4.7
Pitch reset mean values at narrative and non-narrative boundaries

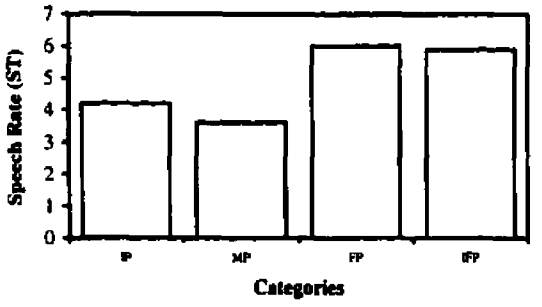


Wider pitch reset values are clearly associated with narrative boundaries, illustrated in the chart above. Statistical analyses yielded significant results (t=-5.924, df=608, p<0.0001). The chart also suggests that while pitch reset at narrative boundaries can be as low as the lowest reset at non-narrative boundaries, the highest values are always associated to narrative boundaries.[52]

Therefore, pitch reset, rather than pitch range, appears to be the most accurate indicator of narrative section boundaries in narrative discourse. It was demonstrated in the previous section that although there is a correspondence between the pitch range of individual intonation units and their position in relation to the narrative sections (see Section 4.5 for a discussion), the effect did not reach statistical significance. Therefore, instead of using pitch range in the assessment, pitch reset values occurring at the end of intonation units, should be employed as the dependent variable, as the statistical significance implies (F (3,609)=12.4156, p<0.0001). Figure 4.8 below displays the mean values for each IU type:[53]

[52] The standard deviation values for narrative and non-narrative boundaries are 4.1 ST and 3.4 ST respectively.
[53] Standard deviations for each case are: 3.8 ST (IP), 3.2 ST (MP), 4.3 ST (FP), and 3.6 ST (IFP).

Figure 4.8
Mean pitch reset values at the end of different intonation units classified according to its position in the narratives



The values in the above chart serve primarily to validate the hypothesis that was already confirmed previously with the analysis of pitch reset in relation to the two different types of boundary in a narrative. As indicated, the reset that occurs subsequent to the last IU in a narrative section (FP) is much wider than those taking place within the narrative sections (MP and IP). This assessment, however, does bring two interesting points to light. Post-hoc analysis (Tukey-Kramer multiple comparisons) demonstrated that intonation units can be classified into two groups, according to the position they occupy in the text: one formed by those IUs that occur solely at the end of a narrative section (FP) and those that occur both at the beginning and at the end as well (IFP), and another formed by the two remaining categories: IP and MP.[54] These results show that what is pertinent is whether the end of an IU corresponds to a major discourse break, rather than whether it occurs at the beginning, middle or end of a major discourse unit. On the other hand, the above results also suggest that those intonation units that simply initiate a new narrative section (IP) are accompanied by a slightly larger pitch reset than those which occur within a narrative section (MP). This is possibly due to the

[54] IP did not differ significantly from IFP, though.

manipulation factor discussed above, where the first intonation unit of a narrative section would become considerably wider or more compressed in terms of pitch range thereby marking the narrative section break by means of a larger pitch reset. As suggested, this manipulation frequently resulted in a larger reset at the end of the first IU of the narrative section, due to the difference in terms of pitch range the first IU in the section normally presented from the following IU.

4.7. Boundary tone as a cue for narrative segmentation

Boundary-marking pitch movements have often been associated with discourse segmentation. In general, melodic contours are said to distinguish between continuation and finality, with rise and level tones indicating the former, and low tones signaling the later (Brown, Currie & Kenworthy 1980; Blaauw 1995; Swerts & Geluykens 1994; Swerts, Geluykens & Terken 1992; Van Donzel 1999; Wichman, House & Rietveld 1997).[55]

The present section investigates whether boundary-marking pitch movements are significantly correlated with narrative structure in the data. Instead of using a distinction between high and low tones, the present study will adopt the terms "low" and "non-low". This is done for methodological reasons discussed above, in Section 4.2.2.[56] Non-low boundaries include both high and mid tones. In accordance with the previous studies in

[55] See, however, Swerts & Geluykens (1994), who suggested that boundary tones can in fact indicate more than just continuation and finality. See also Swerts, Collier & Terken (1994), who made use of a more comprehensive classification of melodic contours in their analysis. Furthermore, it should be noted that a possible complication in such a binary analysis is the concurrence of a rising boundary tone with a pitch accent in monosyllabic words or words with final lexical stress that occurs at the end of an utterance, resulting in the rise of the pitch accent being followed by a rising tone instead of a fall. This phenomenon is known as "accent lending boundary tone" (see Swerts 1994 for more details on this issue).

[56] Refer to footnote 32.

this respect, it is hypothesized that low tones will most typically occur in the final position of a narrative section, and conversely, non-low tones will be generally found within such sections.

The mean values resulting from the consideration of all the narratives in the data do not reveal how boundary-marking pitch movements are distributed within a narrative. In order to illustrate this point, narrative 12 is analyzed below (Figure 4.9), by taking into account the boundary tones occurring at the end of each intonation unit. The classification is displayed linearly, and broken down by narrative sections.

Figure 4.9
Boundary tones (L – low; NL – non-low) at the end of intonation units of narrative 12

Abstract	NL	NL	NL									
Orientation	NL	NL	NL	NL	NL	L	NL	NL	NL	NL	NL	NL
Complication	L											
Evaluation	NL	L										
Resolution	NL	L										
Evaluation	NL	NL	NL	L								
Coda	NL	NL										

Narrative 12 has a total of 26 intonation units. Most of them (21, or 81%) end in a non-low boundary tone. The low boundary tone occurs almost exclusively at the end of narrative sections, i.e., at narrative boundaries (the only exception being the low boundary tone that occurs in the middle of the orientation section). This does not mean, however, that all narrative sections in this story end in a low tone. In the above example, three sections end in a non-low boundary tone (the abstract, the orientation and the coda). In general, 57% of the narrative boundaries in narrative 12 are accompanied by a low boundary tone, whereas 94% of the non-narrative boundaries (i.e., those IU boundaries

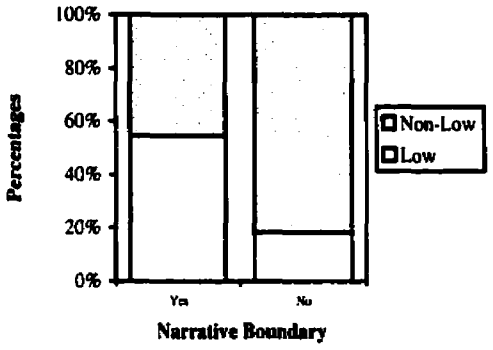
that do not correspond to a narrative boundary) end in a non-low boundary. To a large extent, these results mirror those for the other narratives in the data.

In order to verify whether there is a significant correlation between the type of boundary and the boundary tone, all of the 627 intonation units that make up the narratives in the data were computed. Of these, 468 (75%) ended in a non-low boundary tone.[57] The predominance of non-low tones at the end of intonation units in narratives has already been observed in previous research. Brazil (1997: 93), for example, interprets this as a strategy the narrator employs in order to control the communicative event. Van Donzel (1999: 85) argues in the same way. According to her study, speakers often emply a non-low tone "just to make sure that the listener keeps following the story, or simply as a form of feedback ('do you understand what I'm saying?')".

That boundary tones are employed in the narratives as a function of the different boundary types (narrative and non-narrative) suggests that there is indeed a strong correspondence between the narrative structure and the use of low versus non-low boundary-marking pitch movements. Figure 4.10 below displays the difference (in percentages) in the distribution of low and non-low boundary tones between narrative and non-narrative boundaries. From these results, it may be deduced that narrative boundaries are most of the time (55% of all cases) accompanied by low boundary tones; non-narrative boundaries, on the other hand, systematically end in non-low boundary tone (82% of all cases). This difference is significant (χ^2=59.506, df=627, p<0.0001).

[57] Statistical analysis didn't show any difference in terms of the distribution of low and non-low tones both across narratives and as a function of the different participants. The distribution of non-low tones across narratives ranged from 61% (narrative 01) to 88% (narrative 09); across the different participants, it ranged from 68% (participant 05) to 84% (participant 02).

Figure 4.10
Percentage of distribution of low and non-low boundary tones between narrative and non-narrative boundaries in the data



Much like narrative 12 (see Figure 4.9), the use of low boundary tones in the data appears to be restricted almost exclusively to narrative boundaries. Again, this is not to say that all boundary tones are accompanied by low boundary tones. In fact, the chances for a non-low boundary tone to occur at the end of a narrative section are almost as high as for a low boundary tone.[58] What the results in Figure 4.10 suggest is that whenever an intonation unit ends in a low boundary tone, the chances of that unit coinciding with the end of a narrative section are much greater.

4.8. Acoustic correlates of pitch phenomena

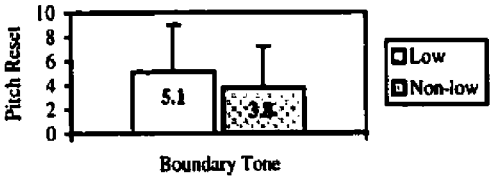
The previous chapters indicated that some temporal phenomena co-occur with statistically significant accuracy with a few intonational variables under investigation in the present research. For example, it was observed that the occurrence of a pause at the end of an intonation unit is closely associated with a higher pitch reset and low boundary

[58] All the narrative boundaries in both narratives 01 and 02, for example, are accompanied by a low boundary tone. In narrative 11, on the other hand, this is true for 16% of the cases only.

tone. Furthermore, it was also noted that both pause duration and speech rate correlate with different boundary tone types. On the one hand, longer pauses are frequently accompanied by low boundary tones; on the other hand, intonation units produced at a higher speech rate frequently end in a non-low tone. It will be investigated whether the intonational parameters that are being considered in this study (pitch range, pitch reset and boundary tone) interact with each other, making it possible to predict the occurrence of one variable, on the basis of the occurrence of another.

First, boundary tone will be correlated with pitch reset. If it is true that both phenomena are employed for indicating narrative section boundaries, it is expected that higher pitch reset will co-occurs with low boundary tones. Figure 4.11 below presents the mean values of pitch reset as a function of its co-occurrence with low and non-low boundary tones:

Figure 4.11
Pitch reset value (in ST) as a function of boundary tone

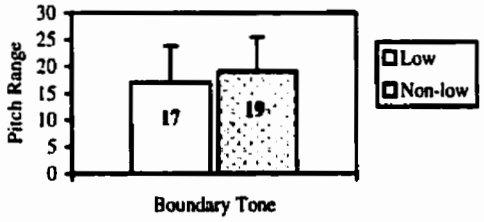


As predicted, larger pitch reset values systematically coincide with low boundary tones. This correlation yielded a significant effect ($t=3.678$, $df=608$, $p<0.0003$), suggesting that the co-occurrence of both wider pitch reset and low boundary tones is better than chance.

Finally, pitch range values were correlated with the boundary tones of each intonation unit in the data. Research on these correlates of intonation in discourse

structure has established that pitch range tends to be higher in utterances located at the beginning of a global discourse units, and lower at the last utterance of such units (Grosz & Hirschberg 1992; Hirschberg & Grosz 1992; Brown, Currie & Kenworthy 1980; Silverman 1987; Avesani & Vayra 1988; Ayers 1992). Even though no statistically significant effect was found in the present study that would confirm the above correlation in the narratives under investigation, a tendency in this direction is definitely observed. Consequently, on the basis of the statistically significant correlation between narrative boundary and low boundary tones, it is expected that low boundaries coincide with more compressed pitch range. Figure 4.12 below displays the means and standard deviation of pitch range as a function of different boundary tones.

Figure 4.12
Pitch range value (in ST) as a function of boundary tone



Although the distance between the mean values is very small (2 ST only), the difference between the two conditions is significant ($t=-3.805$, $df=625$, $p<0.0002$).

These results obviously suggest that certain prosodic features tend to occur simultaneously, and most of them carry a specific (textual) function. In the present case, they signal the existence of a boundary that segments different larger-scale sections in a narrative. This explains, for example, why certain features that were studied here and did not demonstrate to be an accurate narrative section boundary indicator (such as speech range reset), do not co-occur with statistically significant frequency with any other

features that can correctly predict the presence of a narrative boundary. The distribution of all prosodic events realized at narrative section boundaries in the data will be closely examined in the following chapter.

4.9. General discussion and conclusions

The focus of the present chapter was on investigating pitch phenomena as a means of indicating the structure of narrative discourse. The major issues underlying the study presented in this chapter are listed below.

(1) Do pitch range, pitch range reset, and boundary marking pitch movements contribute to the task of making the structure of spontaneous, non-elicited narrative texts more transparent?

(2) If so, how exactly are each of these features related to the way information is structured in narratives?

The three phenomena under investigation (pitch range, pitch range reset, and boundary marking pitch movements) were analyzed in the data both as a function of the information content of the individual narrative sections (discriminated according to the Labovian model of narrative analysis). The same three phenomena were also analyzed as a function of major transition points produced by the segmentation of the narratives into the above-mentioned sections (narrative boundaries).

It was hypothesized, based on previous works dealing with the role of intonation in the broad category of information status, that pitch range values are related to the information content of the narrative sections. Since higher pitch range values are often

associated with the degree of importance certain information has in relation to others in a given discourse, the following hypothesis was tested: complication and evaluation sections are often closely associated with the point of the story, carry the most central elements, and are uttered in a higher pitch.

In terms of discourse segmentation, it was hypothesized that pitch range reset, i.e., the difference in terms of pitch range between two adjacent intonation units, and boundary marking pitch movements, are both very important devices used for the signaling of major boundaries that occur between the various semantically individualized sections in a narrative.

The main findings of this chapter can be summarized as follows.

A clearly identifiable relationship between pitch range and informational content in the narratives of the present study has been observed through the examination of pitch range behavior on the global level. Contrary to the prosodic temporal variables previously investigated, the linear analysis of pitch phenomena over entire narrative texts did not show the existence of a global variation resulting in a cycle that would reflect their structure. Rather than being a function of a linear segmentation, the variation in pitch range in the present data appears to be related more to the content that each individual narrative section conveys.

As predicted, the sections that are often considered to be essential in a narrative (complication and evaluation) were usually uttered at a higher pitch range. The results also indicated that coda sections are, in general, similar to complications and evaluations in terms of the use of pitch range. It was suggested that the reason for this result could be

explained by taking into account the fact that the codas in the narratives contain much material that is intrinsically evaluative.[59]

Some studies on the role of pitch phenomena in narrative discourse have already observed that pitch has a decisive influence on the structuring of stories. Kumpf (1987), for example, suggested, on the basis of her analysis of elicited narratives told in English by ESL speakers that events and non-events in a narrative can be distinguished in storytelling by means of pitch strategies. According to her findings, speakers often use a "normal", or "declarative" intonation pattern when they are reporting events; asides, incidental events, and non-essential explanations are downtoned, signifing a lack of "importance" or "focus". Non-events, such as "expressions of surprise" and "evaluations" and events that are considered "central" in a story are, on the other hand, uttered in high pitch. Pakosz & Flashner (1988) argue in a similar fashion, noting that pitch, among other prosodic features, is systematically employed to separate foreground from background information in storytelling.

Instead of employing a (closed) binary classification of narratives into utterances that report the sequence of events in a story and statements that serve only to support the point of the narrative, the present investigation opted for a more detailed classification of the narrative structure. Narratives were segmented into (possibly) six individual sections, characterized by the informational content carried out by a group of intonation units. Each section was then examined for a potential global pitch range value in order to set one section apart from another. Statistical analysis showed significant differences between two groups only. One comprised the complication, the evaluation, the resolution

[59] It must be stressed here again that a section is defined in terms of what it represents as a whole. The fact that a coda contains a great amount of evaluative material does not necessarily mean that it would rather be considered an evaluation instead.

and the coda sections, characterized by a higher pitch range value; and the other comprehended the abstract and the orientation section, characterized by their lower pitch range values.

Although these results may be interpreted as a reflection of the above-mentioned binary classification because they suggest that the information can be distinguished in terms of relevancy, no distinction between event/non-event or sequence/non-sequence[60] in terms of global pitch range could be observed here. Events and non-events were grouped together into a major group conveying central information in the narrative. This suggests that this is actually the most relevant attribute for the classification of information in terms of pitch range values.

Further evidence for this was also observed at the local level, by examining individual intonation units in the narratives. Some utterances, although occurring neither within the complication/resolution section(s), nor within the evaluation section(s) of a narrative, were accompanied by a high pitch range. In most cases, such utterances contained a considerable amount of evaluative features, resulting in the use of the higher pitch range values. This was specifically the case for coda sections, as stated above.

It was also suggested that speakers sometimes employ a narrow pitch range for evaluative purposes, but not to indicate that the information in question is of little or no importance. This implies that contrastiveness, rather than actual values, is what appears to be relevant here.[61]

---

[60] See Hopper 1979, Givón 1982 and Hopper & Thompson 1956 for further details on these concepts, as related to grammatical structures in discourse, specifically in narrative discourse.

[61] In some specific cases, where (a) intonation unit(s) was (were) uttered in a compressed pitch range within a narrative section that was supposed to present an overall high pitch range value, it was observed that speakers attempt to compensate for the low value(s) by increasing the pitch range of the other intonation unit(s) in the section.

The findings of the present chapter not only corroborate the assumption that pitch variation in discourse is closely related to semantic and pragmatic factors, but specifically shows that from the perspective of global pitch behavior the segmentation of narratives into semantically determined topical constituents (narrative sections), based on an independently-motivated theory of narrative analysis (the Labovian evaluative model) is a valid one.

The other two features, pitch range reset and boundary-marking pitch movement, were examined in terms of their role in the segmentation of narrative on the local level, by way of their signaling of narrative section boundaries.

It was demonstrated that there is a tendency for pitch range values of individual intonation units to vary according to the position of the intonation unit within a narrative. More expanded pitch range values are commonly associated with intonation units that begin a narrative section; more compressed pitch range values are, on the other hand, generally associated with the last intonation units of a narrative section. This finding is not new in the study of prosodic characteristics of discourse boundaries; it merely replicates the results of previous research (Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Brown, Currie & Kenworthy 1980; Silverman 1987; Avesani & Vayra 1988; Ayers 1992). However, since the variation in pitch range values, as a function of intonation unit position, did not show any statistically significant effect here, it also suggestes that the difference of pitch range values of adjacent intonation units, rather than the pitch range values themselves, is relevant to the marking of narrative section boundaries.

The examination of pitch range reset as a cue for narrative segmentation revealed that this feature is a highly significant one: larger-scale information units in narratives

(narrative sections) tend to be phonetically encoded by means of a higher pitch reset value. Low values, on the other hand, generally signify topical continuity. Based on previous findings that considered the melodic closeness of intonation units as properties of individual narrative sections, a larger difference between pitch range values at narrative boundaries was to be expected. This is to say that it would reflect the differences of pitch range values of narrative sections on the global level. However, it was also shown that narrative sections only differ in terms of global melodic patterns in basically two ways: by displaying either a compressed or an expanded pitch range. It is very unlikely that a section characterized by a compressed range is systematically followed by a section displaying an expanded range, and so on. Furthermore, a considerably amount of cases displaying pitch range value disparities within a narrative section, causing pitch reset to occur at places other than narrative boundaries, could also be observed. It was suggested that, in some cases, pitch range values were manipulated in order to force a stronger reset at a narrative boundary. As a result, high pitch reset values also occurred within a narrative section. It appears thus that a strong reset occurs at a narrative boundary, even if that results in the occurrence of a strong reset elsewhere. This clearly indicates that this prosodic feature is considered, from the point of view of the speaker, as a very important cue for narrative segmentation.

Boundary-marking pitch movements, it was found, in agreement with the literature, that speakers tend to restrict the use of low boundary tones to mark the end of a global topic unit. In other words, intonation units belonging to a single narrative section are often accompanied by a non-low boundary tone, indicating continuity. This is not to say, however, that non-low boundary tones do not coincide with narrative section

boundaries. They do, as a matter of fact, in almost fifty percent of the cases. So, if on the one hand, speakers do not systematically signal the end of a narrative section with a low boundary tone, suggesting that this is not a very recurrent cue for narrative segmentation. If, however, on the other hand, they seem to be aware of the implicit meaning of finality conveyed by the use of low boundary tones, this feature become a very accurate discourse boundary predictor.

Finally, this chapter also demonstrated that the pitch features under investigation co-occur not only with other prosodic temporal phenomena, as demonstrated in the previous chapters, but also with each other in a very systematic and predictable way. For example, boundary tones were demonstrated to interact significantly with both pitch reset and pitch range: low boundary tones are often accompanied by a stronger pitch reset and a more compressed pitch range. Further investigation on how the various prosodic variables analyzed here interact with each other, and which features are more important for the segmentation task will be addressed in the next chapter.

It has been shown here that pitch phenomena, specifically pitch range, pitch reset, and boundary tones, contribute significantly to the clarification of the topical make-up of a narrative, by separating successive narrative sections and characterizing all of them in terms of global pitch range values. The hypotheses concerning the use of pitch phenomena in narrative discourse were all confirmed, unlike those put forward for the temporal variables, suggesting that pitch is the most reliable indicator of narrative structure.

The next chapter will bring together the results from chapters 2, 3, and 4 in a general discussion of the role of prosodic features in narrative segmentation. It will also present the final conclusions as well as suggestions for future research.

General Discussion and Conclusions

I like the gardens with good bones and affirmed underlying structure, well-marked paths, well-built walls, well-defined changes in level.

Russell Page,
The Education of a Gardener

Abstract

This concluding chapter provides a general discussion of the findings reported in the present study. An outline of each chapter will be given in order to account for the contribution of the present work to the study of the communicative (demarcative) function of prosody in storytelling. The general conclusion is that narrative discourse presents a reasonably predictable structure, which is systematically signaled by different prosodic cues. The findings are summarized, followed by the acknowledgement of the limitations in this study and future research is suggested.

5.1 Introduction

The present investigation has dealt with the prosodic features that are frequently employed in spontaneous speech as structural devices. An empirical model specifically designed for narrative discourse is used to demonstrate that pause, speech rate, and pitch phenomena occur systematically in storytelling as cues to narrative structure.

The narratives used in this experiment were extracted from "spontaneous interviews."[1] Recordings were made in a sound proof laboratory in accordance with the standards required for acoustic analysis. All of the participants were native speakers of Brazilian Portuguese and all participated in this research on a voluntary basis. Although summoned to "talk freely" on any of the topics indicated in a list presented to them, the participants were not instructed to tell stories. The selected narratives appeared spontaneously in their talk. It was assumed that these narratives reflected the speech style that is often referred to as "spontaneous." The factual spontaneity of the data is, however, a matter of debate, since the material is not ecologically valid, from a methodological standpoint.[2] Authentic spontaneous speech can only be found in unobserved settings, under natural and legitimate conditions. Obviously, such an ideal situation makes the task of data collection nearly impossible (hence the "observer's paradox").[3]

In order to relate the prosodic features to the structure of the narratives in the data, the Labovian Evaluative Model (Labov 1972) was chosen as the framework for the present analysis. A pilot experiment demonstrated that this model, which was based on the examination of oral narratives, is reproducible. Seven trained labelers were able to determine the boundaries of all the narratives in the corpus and to correctly categorize the narrative sections in a better than chance rate by taking into account the informational content of the narratives only.

Supported by this finding, the narratives in the data were segmented and labeled by two discourse analysis experts using the Labovian Model as the theoretic framework. Acoustical analyses were used in order to examine whether the prosodic variables under

---

[1] Refer to Chapter 1, Section 1.5.1 for a discussion about the concept of "spontaneous interviews."
[2] See relevant comment on Chapter 1, Section 1.5.2.
[3] See discussion on Chapter 1, Section 1.5.1.

investigation systematically correlated with boundaries and narrative sections, as established by the Labovian Model. It should be noted that the purpose of this study was not to validate the Labovian Model of narrative analysis, but to demonstrate that prosody can give crucial insight into the way narratives are organized, regardless of which model is being employed.

Three variables were selected for the analysis: pause, speech rate and pitch range. It was hypothesized that these variables have a decisive role in the segmentation of narratives at both the global and the local levels. At the local level, it was hypothesized that the boundaries occurring between narrative sections are prosodically different from other boundaries because they are often accompanied by longer pauses, higher pitch reset, and low boundary tones. At the global level, it was hypothesized that narrative sections may be characterized by the recurrence of specific prosodic phenomena, such as a greater occurrence of (longer) pauses at sections that are primarily evaluative, an acceleration of speech rate at sections that are characterized by their relatively low relevance, and the use of an overall higher pitch range in sections that are considered to be central in the story.

The thesis was structured in order to consider each prosodic variable individually. First, the contribution of pause to the structure of narrative text was investigated. The resulting analysis indicated that pausing is consistently used in narratives as a means of segmenting "chunks of information" that often correspond to the concept of "narrative sections." Specifically, it is not only that pauses occur more often at the end of narrative sections, but also that most of these pauses are consistently longer than the pauses occurring elsewhere. Neither pause occurrence nor pause duration were found to be

reliable cues for the characterization of narrative sections, however. Speakers do not seem to use pausing as a strategy to identify the semantic properties of different parts of a story. Rather, pausing is consistently used to make the segmentation of narrative texts into "sections" characterized by a semantic consistency transparent to the audience. This could be observed through the incidence of a pattern of varying fluency governed by pause occurrence and pause duration matching the segmentation of narratives into sections.

Following the analysis of pauses in the data, speech rate was investigated as a possible acoustic prosodic cue to narrative structure. Similar to what was observed in connection with pausal phenomena, a cycle displaying varied speech rate measurements was also found. This pattern, too, reflected the way narratives are segmented, confirming that prosodic temporal features in speech, as a reflection of the cognitive process of planning and execution, play a very significant role in the organization of narrative texts. In relation to the characterization of narrative sections, there is a tendency for speech rate and textual relevancy to be closely associated. This finding is in accordance with Hypothesis 3 (refer to Chapter 1, Section 1.4.2). Although statistical analyses did not show any significant effect in this direction, the existence of this tendency must be acknowledged.

Finally, pitch phenomena were also considered. In contrast to the temporal variables previously investigated, pitch did not render a cycle of low and high values, suggesting that this cyclical phenomenon might be a characteristic restricted to the temporal property of speech. Pitch varied as a function of the informational content of the individual narrative sections. Higher pitch range was associated with sections that are

often considered to be essential in a story. Although this might suggest that individual narrative sections can be characterized by distinctive pitch range values, statistical analyses demonstrated that narrative sections can be classified into three different groups only, according to the size of the pitch range. Concerning the role of pitch in the segmentation of narratives, it was found that the difference of pitch range values of adjacent intonation units (pitch reset) is a very significant cue — narrative sections tend to be separated from one another by means of a higher pitch reset value. Furthermore, boundary-marking pitch movements were also found to be associated with narrative structure, because they often indicate whether a given narrative section has come to an end or not. Low boundary tones are regularly used to mark the end of a narrative section.

The present chapter is organized as follows: the introductory section wraps up an overview of the study, with a summary of the main findings in each chapter. Section 5.2 proceeds with a general discussion of these findings. The role of pause strategies, speech rate, and pitch phenomena will be considered on both the local level (Section 5.2.1) and the global level (Section 5.2.2), in order to test the validity of the hypotheses that were initially proposed (Chapter 1, Sections 1.4.1 and 1.4.2). The chapter concludes by identifying limitations of the present work, stating the practical and theoretical contributions of the study to different fields of research, and presenting some suggestions for future research.

5.2 Prosodic features as cues for narrative structure

As previously stated, the prosodic variables under investigation in the present dissertation were analyzed on two different levels. The four working hypotheses were formulated by taking into consideration these separate levels (see Chapter 1, Sections 1.4.1 and 1.4.2). Hypotheses 1 and 2 were formulated in connection with the function of the variables at the local level, with regard to their role at different boundaries. Hypotheses 3 and 4 were formulated in connection with the function of the same variables at the global level, with regard to their possible importance for the characterization of different narrative sections. The present section will provide a summary of all the findings by considering these two levels of analysis. The purpose of this summary is to compare all the variables examined, in order to understand how each of them contributes to the segmentation and the semantic characterization of narrative texts.

5.2.1 Prosodic features at the local level: marking narrative boundaries

Assuming that narratives are composed of small sections that convey a coherent message, it was hypothesized that prosodic cues are recurrent and efficient for indicating the boundaries of such sections by narrators. The occurrence of these boundaries would provide evidence that narrators are aware of the underlying structure of narrative texts. In order to test this hypothesis, narrative section boundaries were compared with boundaries that take place between intonation units. In particular, it was hypothesized that:

1. Longer pauses tend to occur more often at narrative boundaries than elsewhere;
2. Pitch reset, or the difference in terms of pitch range between two adjacent intonation units, is higher at narrative boundaries than at intonation unit boundaries; and
3. Low boundary tones tend to occur at the end of narrative boundaries, whereas non-low boundary tones are generally found within narrative sections.

All three hypotheses were confirmed by the analysis of the present research. Pause and pitch phenomena are reliable cues to narrative segmentation at the local level, as illustrated in Table 5.1 below.

Table 5.1
Distribution of the prosodic variables under investigation as a function of the type of boundary: those that occur at the end of a narrative section (Yes) and those that occur at the end of an intonation unit (No). Pause occurrence is expressed in terms of percentage; pause duration, in terms of seconds; pitch rest, in terms of semitones; and low boundary tone occurrence, in terms of percentage

Narrative Boundary	*Pause Occurrence*	*Pause Duration*	*Pitch Reset*	*Low Boundary Tone Occurrence*
Yes	65 %	0.92 sec	5.9 st	55 %
No	43 %	0.74 sec	3.8 st	18 %

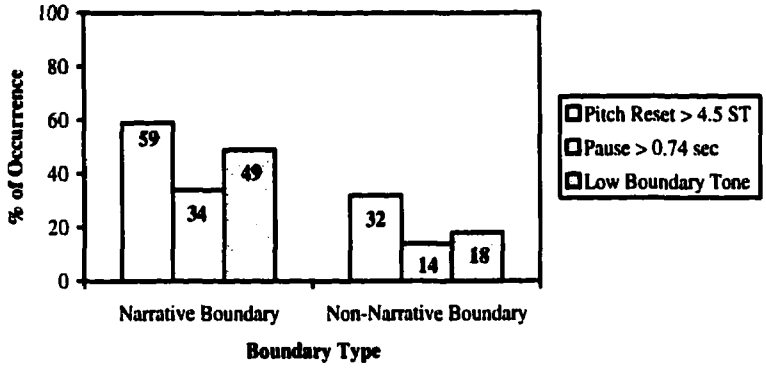
The two conditions (narrative vs. non-narrative boundary) differ from each other in all cases. It should be pointed out, however, that none of these variables, when considered in isolation, constitute sufficient evidence for the existence of a narrative section boundary. This is to say that the occurrence of a (long) pause, for example, does not automatically signify the end of a narrative section. Long pauses, as discussed in Section 2.8, Chapter 2, can be used for other purposes in a narrative text. A high pitch reset does not necessarily imply the existence of a narrative boundary either, and the same can be said for the occurrence of a low boundary tone. What the results suggest is

that although narrative boundaries cannot be identified by the occurrence of any of these variables taken in isolation, the chances of these variables to occur in such a pattern are greater when the position where they take place coincides with a narrative boundary.

Throughout the analysis, it was demonstrated that some features systematically co-occur with others. For example, pause occurrence is closely associated with higher pitch reset values and low boundary tones; larger pitch reset values analytically coincide with low boundary tones; low boundary tones are connected with longer pauses. These findings not only validate the hypothesis that the above-mentioned prosodic features do occur as narrative boundary markers, but also suggest that a difference in terms of "strength" may exist and be effective.

Hypothesis 2 (see Section 1.4.1) states that the greater the number of prosodic cues associated with a boundary, the higher the chances of its being a narrative section boundary. In order to test whether this is the case, the frequency of pause, pitch reset and boundary tone occurrences were first computed, then the possible combinations of these variables were compared. Figure 5.1 below presents the distribution of all the three variables at narrative and non-narrative boundaries. For this analysis, only pauses longer than 0.74 sec and pitch reset values higher than 4.5 ST were considered. These values correspond to the average of both variables, computed by taking into account narrative and non-narrative boundaries.

Figure 5.1
Frequency of occurrence of high pitch reset, long pauses and low boundary tones at narrative and non-narrative boundaries



The numbers in Figure 5.1 clearly demonstrate that these three variables are systematically employed at narrative boundaries more often than at non-narrative boundaries. The distribution of occurrence, however, is very similar in both cases. High pitch reset values occur with a relatively higher frequency at both narrative and non-narrative boundaries, while long pauses occur with a relatively lesser frequency in both cases. These results suggest a primacy of pitch reset at marking narrative boundaries over the other variables, if the same pattern did not occur at non-narrative boundaries as well. Therefore, rather than being a specific property of narrative boundaries, the higher recurrence of elevated pitch reset values should be regarded as a characteristic of intonation unit boundaries in general. Figures 5.2 and 5.3 below bring more detailed information to light on the frequency of distribution of high pitch reset, long pauses, and low boundary tones in both narrative and non-narrative boundaries, broken down by the different participants.

Figure 5.2
Frequency of occurrence of high pitch reset, long pauses and low boundary tones at *narrative boundaries*, broken down by participants

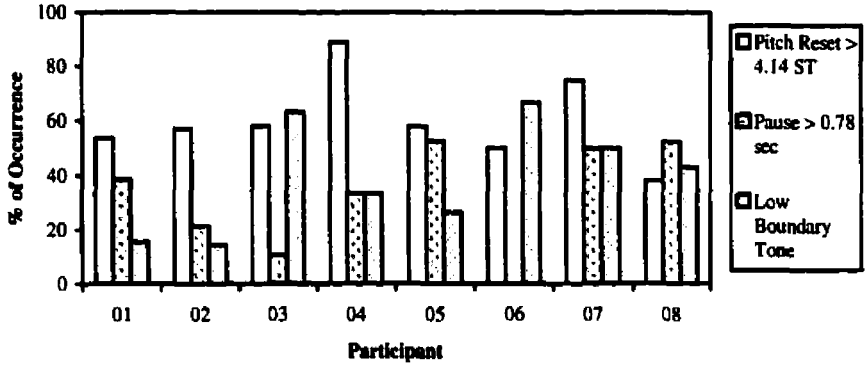
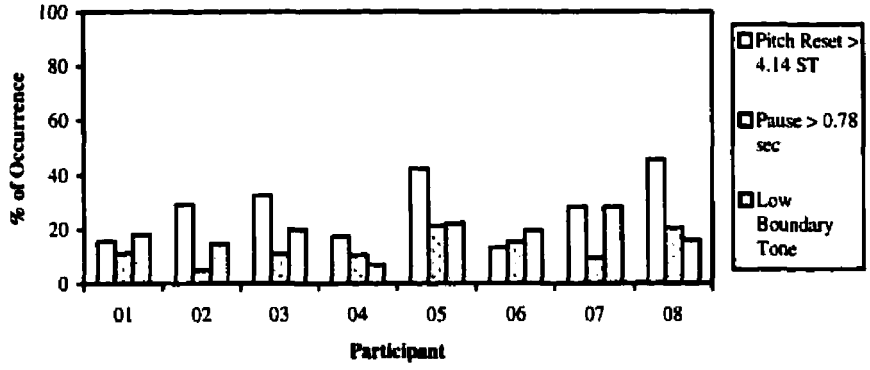


Figure 5.3
Frequency of occurrence of high pitch reset, long pauses and low boundary tones at *non-narrative boundaries*, broken down by participants



Figures 5.2 and 5.3 demonstrate that the participants made different use of the prosodic cues at both narrative and non-narrative boundaries. For example, while participant 08 used long pauses more often than any other cue to mark a narrative boundary, participant 06 did not make use of long pauses at all at this position. Of all the participants, 04 made more extensive use of high pitch reset at narrative boundaries. Participant 06 used low boundary tones at narrative boundaries more often than any other. In general, the pattern in Figure 5.1 could not be replicated by taking the

distribution of the phonetic cues as a function of the individual participants. The use of long pauses and low boundary tones vary greatly from participant to participant. However, high pitch reset seems to be the most systematically employed feature in all cases for the majority of the participants.

In order to answer the specific question of whether the number of prosodic features at a given boundary may, *per se*, function as a cue for its characterization as a specific boundary type, a quantitative analysis of all possible combinations of the three phonetic variables under investigation was carried out. Figure 5.4 below displays the result of this analysis. In the graphic, the combinations are expressed in a binary fashion, marking either the presence of the variable (1), or its absence (0). The combinations are interpreted taking into consideration the order of the variables, as coded by the binary classification. The arbitrary order used here is as follows: high pitch reset + long pause + low boundary tone. Following this, 011, for example, should be interpreted as the absence of a high pitch reset and the presence of both a long pause and a low boundary tone; 010, as the absence of a high pitch reset and a low boundary tone, but the presence of a long pause; and so on.

Figure 5.4
Frequency of occurrence of the three phonetic cues combined at narrative boundaries

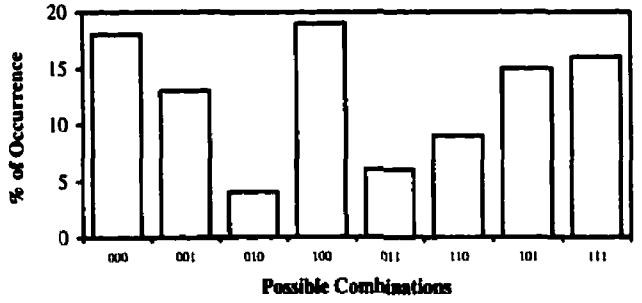
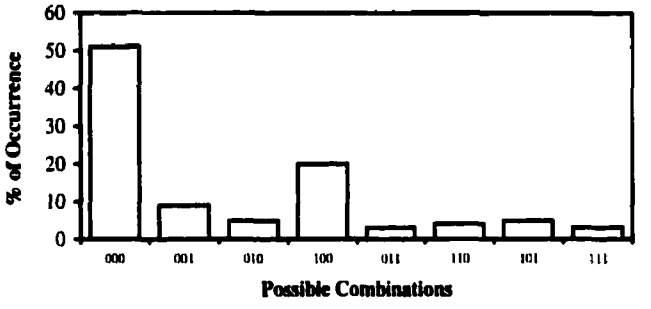


Figure 5.4 shows that it is as common for narrative section boundaries to be accompanied by all three variables at the same time (16%), as it is for them to be followed by no cue at all (18%). However, the percentage of instances when at least one of the cues is present at this position (all cases, except 000) is statistically significant (82%). The combinations of two cues (011, 110, and 101) cover 30% of the cases, while the combinations that involve the presence of just one variable (001, 010, and 100) correspond to 36% of the cases. These numbers are the result of the greater frequency of occurrence of high pitch reset values, as shown in Figures 5.1, 5.2, and 5.3. The combination 100, which means the presence of only a high pitch reset, is the most recurrent one at narrative boundaries (19%). High pitch reset values also have a significant influence on combinations 101 and 110, which are the most frequent among the combinations that are characterized by the presence of two variables. Following pitch reset values, low boundary tones are also quite frequent at narrative boundaries, as Figure 5.4 reveals. Combination 001 is one of the most recurrent, with 13% of the cases. This frequency is slightly higher when boundary tone is combined with pitch reset, reaching 15% of the cases. The results in Figure 5.4 corroborate the findings displayed in Figure 5.1 in relation to the low frequency of long pauses at narrative boundaries, as compared to the other variables.

Evidently, the results in Figure 5.4 have meaningful interpretation when compared with the frequency of the same combinations at non-narrative boundaries. If the results are similar in terms of distribution, then the considerations made above are not to be regarded as a characteristic of narrative section boundaries only.

The frequency of occurrence of the eight possible phonetic cue combinations at non-narrative boundaries is presented in Figure 5.5 below. This figure clearly reveals that the phonetic cue combinations have a very different distribution at non-narrative boundaries, when compared with their distribution at narrative boundaries.

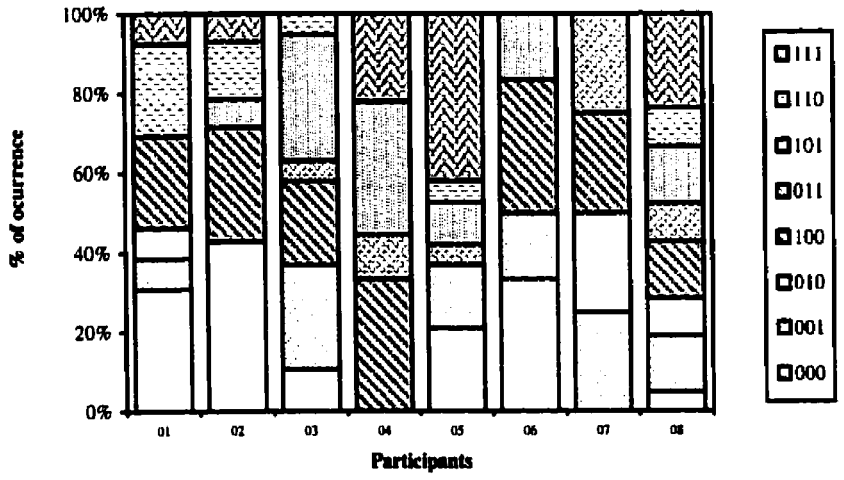
Figure 5.5
Frequency of occurrence of the three phonetic cues combined at non-narrative boundaries



In contrast with what happens at narrative boundaries, the total absence of phonetic cues at non-narrative boundaries (combination 000) corresponds to more than 50% of the cases, while the presence of all the phonetic cues at the same time (combination 111) is one of the less frequent combinations (3%). The only combination that occurs in more than 10% of the cases is, not surprisingly, 100 (high pitch reset only). Therefore, the pattern of distribution of possible phonetic cue combinations differs greatly from narrative boundaries to non-narrative boundaries. However, as Figures 5.4 and 5.5 suggest, it is not the presence of a greater number of cues that increases the chances of a boundary being characterized as a narrative boundary. This observation confirms what was proposed by Hypothesis 2 (Chapter 1, Section 1.4.2), that it is the absence of these combinations that typifies non-narrative boundaries.

Once again, it is important to consider the distribution of the possible phonetic cue combinations as a function of the different participants in the experiment. Figures 5.6 and 5.7 below display detailed information on this distribution. Figures 5.6 and 5.7 illustrate the different possible combinations that the participants use at narrative and non-narrative boundaries. For example, contrary to the trend in Figure 5.4, participant 07 did not make use of combinations 000, 010, 011, 110, and 111 at narrative boundaries at all. Moreover, all four combinations employed by participant 07 were distributed equally. Participant 05 marked narrative boundaries in 42% of the cases with all three prosodic features, while participants 03 and 07 never did so. Participant 02 marked narrative boundaries in 42% of the cases with the lack of all three prosodic variables, while participant 04 always marked a narrative boundary with at least one cue.

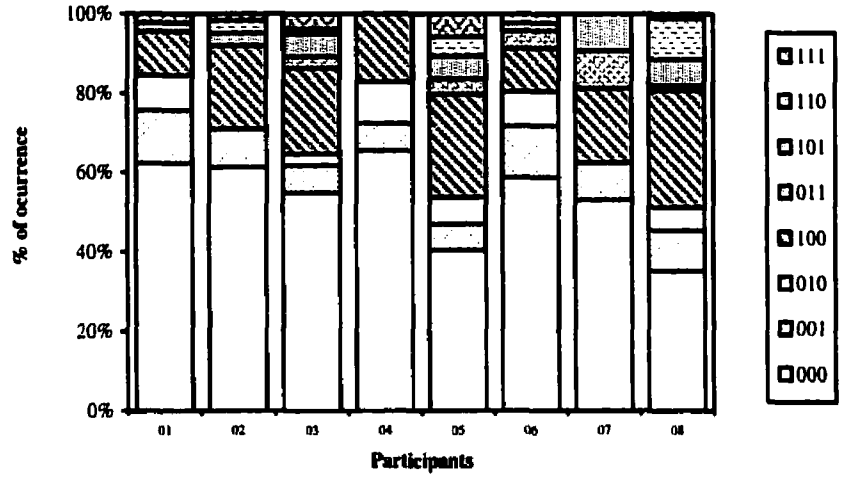
Figure 5.6
Frequency of distribution of the eight possible phonetic cues combinations at narrative boundaries, broken down by participants



The distribution of the possible combinations at non-narrative boundaries parallels the trend established in Figure 5.5. At this position, combination 000 (no prosodic cues) was used in more than 50% of the cases, for most participants (the

exceptions being participants 05 and 08). All participants employed combination 100 (high pitch reset only) systematically, while combinations 011, 110, and 111 were only sporadically used.

Figure 5.7
Frequency of distribution of the eight possible phonetic cues combinations at non-narrative boundaries, broken down by participants



The findings reported here can be summarized as follows: evidence for the validation of a subjacent structure of narrative texts could be established by taking into account the prosodic information contained in spontaneously produced oral narratives. Looking at it from another perspective, this finding suggests that speakers make regular use of prosodic cues to evince the segmentation of narrative texts into "chunks of information" that closely correspond to the notion of "narrative section" proposed by Labov (1972). Long silent pauses, low boundary tones and high pitch reset values are systematically employed at intonation unit boundaries that coincide with a narrative section boundary, as markers of these major discourse breaks. Consequently, they indicate the structure of the narrative. The way these cues are employed vary from speaker to speaker. However, in general, the fewer the number of the above-mentioned

cues are used at a boundary, the lesser the chances are of it being a narrative section boundary.

5.2.2 Prosodic features at the global level: characterizing narrative sections

Prosodic features were also analyzed at the global level, with the purpose of identifying possible patterns that serve to characterize different narrative sections. It was hypothesized that if semantically individualized narrative sections present an identifiable prosodic paradigm, further evidence for the existence of an underlying narrative structure would be given (refer to Hypotheses 3 and 4, in Chapter 1, Section 1.4.2). Pause occurrence and pause duration, pitch range and speech rate, were investigated. It was hypothesized that:

1. Pauses tend to occur more often and tend to be longer at sections that are characterized by the complex process of interpreting, such as evaluations and codas;
2. Faster speech is associated with sections that are characterized by relatively low relevance to the narrative, such as orientations and abstracts. Highly relevant sections in a story, such as complicating action and evaluation are characterized by a slower speech rate;
3. Pitch range interacts with the degree of importance of the section in a narrative: sections that are considered highly significant in a story, such as evaluations and complications, for example, are characterized by an overall higher pitch.

The results summarized in Table 5.2 below indicate that, in most cases, prosodic features do not interact with individual narrative sections.

Table 5.2
Distribution of the prosodic variables under investigation as a function of the different narrative sections. Pause occurrence is expressed as a ratio of number of potential pause position; pause duration is given in seconds; speech rate (syllables per second), and pitch range is presented in semitones. Standard deviation values are given in parentheses

Narrative Section	*Pause Ratio*	*Mean Pause Duration*	*Mean Speech Rate*	*Pitch Range*
Abstract	0.085	0.73 sec (0.34 sec)	6.2 s/s (2.2 s/s)	15.3 st
Orientation	0.115	0.73 sec (0.37 sec)	5.6 s/s (1.1 s/s)	15.9 st
Complication	0.099	0.77 sec (0.40 sec)	5.5 s/s (1.2 s/s)	19.1 st
Evaluation	0.097	0.74 sec (0.46 sec)	5.5 s/s (1.8 s/s)	21 st
Resolution	0.103	0.86 sec (0.61 sec)	5.5 s/s (1.8 s/s)	18 st
Coda	0.039	0.98 sec (0.91 sec)	6.1 s/s (1.6 s/s)	19.7 st

Out of the variables considered in this analysis, only pitch range yielded a significant result.[4] Pause occurrence, pause duration, and speech rate do not vary, in general, as a function of the individual narrative sections. It is important to note, however, that while no significant effect could be found for speech rate, a trend in the hypothesized direction was observed. Complications, evaluations, and resolutions form a group of similar lower speech rate values, while abstracts and codas form another group of relatively higher speech rate values. As for pause, no such trend could be established, suggesting that the content of individual portions in a narrative has little or no influence on pausing strategies.

The results surrounding pitch range indicate that narrative sections differ from one another in terms of global melodic patterns in two ways: by displaying either a compressed, or an expanded pitch range value. Abstracts and orientations are included in the first group; complicating actions, evaluative sections, resolutions and codas, in the

[4] For statistical analysis of each individual variable, refer to the corresponding chapter where they are examined.

latter. This binary classification corroborates the assumption that pitch range is often associated with general textual meanings, rather than ratifying the so-called event/non-event distinction, as proposed in many prosodic analyses of narrative texts.[5]

This leads to the premise of Hypothesis 4 (Chapter 1, Section 1.4.2), which proposes that the content of the narrative may have an effect on the use and distribution of the prosodic variables under investigation. However, it was demonstrated that certain prosodic features may display variable behavior from what it is expected, and that this might be due to the content of either the narrative as a whole, or the particular section in which such an irregularity occurs, or both. For example, the employment of long pauses in storytelling, which are typically used to mark narrative boundaries, may have a stylistic function, as discussed in Chapter 2, Section 2.8. In stories dealing with suspense, as in the case of narrative 01, for instance, long pauses are often used evaluatively, as a strategy to maintain the listener tension.

Furthermore, the relation that the various sections in a narrative have with each other also has a significant effect on the way prosodic features are employed. For example, it was demonstrated that narratives present an "encoding cycle" characterized by the alternation of pause frequency in every other narrative section. However, it was also shown that the more a narrative section is embedded in another in terms of informational content, the less likely they will differ in terms of the pause to speech ratio. This is also observed for speech rate. The variation in speech rate in narrative sections presented a somewhat asymmetrical cycle: the speech rate of the subsequent section tended to follow the upward or downward movement of the previous section. It was

[5] See discussion in Chapter 4, Section 4.9.

suggested that this only happened when the two sections were connected by means of an evaluation.

The informational content of the individual intonation units may also contribute to the prosodic characterization of the narrative sections. The analyses in Section 3.6, Chapter 3, for example, suggested that sometimes the information conveyed at the local level (i.e., at the intonation unit) is of much more importance for the determination of speech rate than that conveyed in a more global discursive level (i.e., at the narrative section). This often resulted in the rate of a given section not displaying the expected value for that specific section, following the trend that was verified for all the narratives in the data. The conclusion was that the primary function of a narrative section does not always determine its corresponding speech rate — the elements within the section should be taken into consideration all the time.

The manipulation of prosodic features as a function of the informational content of the narrative was clearly exemplified with numerous references to the evaluative function of storytelling.[6] It was demonstrated that narrators systematically make use of prosodic devices with evaluative purpose.

One of the most recurrent evaluative features in narratives that requires a certain amount of suspense is the use of long pauses, which have been called "semantic" or "stylistic" pauses. Long pauses are often used in narratives to maintain the tension for the audience and create suspense. They should not be confused with structural pauses, although sometimes they coincide with them. The analysis of narrative 01 in Chapter 2 (Section 2.8) exemplifies this evaluative function of pauses in storytelling.

---

[6] Refer to Chapter 1, Section 1.2.2 for a discussion.

Speech rate is also frequently employed as an evaluative feature. The opposition between fast and slow speech rates commonly serves to separate the essential content of the discourse from what is trivial information, thereby reflecting an intrinsic evaluative function of this temporal variable. The expressive use of speech rate is well exemplified in storytelling. Speakers often use faster speech rates in narratives to emphasize the tension of certain dramatic scenes. Slow rates are often employed to demonstrate contempt, disappointment and/or grief. Additionally, speech rate functions as a mimicry device in narratives. When reporting events that occurred hastily in the story-world, speakers tend to use a faster speech. Slow paced events are reported accordingly, suggesting a parallel pacing of sequence of events to speech rate. It was furthermore suggested that speech rate may be employed as a face-saving maneuver. Attitudes, when uttered in a given speech rate, may carry a subjacent judgment towards themselves. When certain attitudes that are considered to be socially unacceptable are reported by their own authors, the negative impression may be minimized through the manipulation of the speech rate.

Finally, it was demonstrated that pitch range also has a crucial role in signaling what is more or less relevant in the storytelling process. Very often, this information is given at the local level, rather than at the global level. Intonation units that are characterized by their highly evaluative value are frequently uttered in a higher pitch range. Pitch range also characterizes informational content at the global level. Evaluative sections and sections containing a significant amount of evaluative material are commonly uttered in a higher average pitch range. It should be mentioned, however, that depending on the context, low pitch range values may have an evaluative effect as well.

This is specifically the case at sites where pitch range values are supposed to be high. As previously observed (Section 3.1), any element or feature that deviates from an expected behavior can be regarded as evaluative.

The use of these prosodic variables in signaling evaluation in narratives assists the audience in appreciating the narrative more efficiently. It remains to be shown what the impact of such features in perception is and to what extent they are actually used in narratives.

5.3 Limitations, contributions and future directions

The present experimental investigation has demonstrated that storytellers seem to be aware of the existence of an underlying narrative structure, and that this awareness is evidenced through the systematic use of various prosodic devices. It was primarily designed to describe prosodic patterns that occur in the production of this specific discourse genre, as a means of signaling its inherent structure. For that reason, an *acoustic* approach was adopted, rather than a *perceptual* one.

It is well known that any study dealing with the role of prosody in speech should consider both the production and the perception aspects of the prosodic features under investigation. In other words, it is important that such studies demonstrate the extent to which the features that were described in production are relevant in perception. In this sense, the present dissertation is limited, because it does not bring into account the role the prosodic features play from the point of view of perception. Therefore, it is proposed

that a follow-up to this study, testing the perceptual validity of the features here investigated, be undertaken.

It would be particularly interesting to test to what degree the prosodic features that were presently investigated assist in the labelers' judgments of narrative structure. It was demonstrated that when narratives are considered without reference to any acoustic cue, labelers' judgments on narrative structure closely match the Labovian model (see discussion on Chapter 1, Section 1.5.4). Would the presence of acoustic cues interfere with these judgments? It would also be extremely important to know which features, from the perceptual standpoint, are more important for the establishment of the narrative structure. The perceptual analysis of the present data should be considered a natural follow-up to this study.

A very important prosodic feature has been left out of consideration in the present investigation: intensity. Some works have already demonstrated that this feature has a pre-eminent demarcative function in discourse (Grosz & Hirchberg 1992; Hirschberg & Grosz 1992; Grosjean 1983), while others have illustrated how it can be used evaluatively in dialogues (Günthner 1997). The analysis of intensity in spontaneous speech is, however, surrounded by several methodological drawbacks that, most of the time, make the results derived from it of questionable value.

One of the most obvious problems in such a study is the collection of good working data. It is well documented that amplitude varies as a function of the physical distance of the source from the recording device. Because it is nearly impossible to maintain the head position of an interlocutor who is supposed to speak spontaneously for more than 30 minutes in a fixed position, this poses obvious logistical problems. Of

course the researcher can use a head-mounted microphone in order to avoid variation of amplitude in the recording process. Although this is a good working solution for a small and brief task, it might become uncomfortable for a more prolonged one, especially when the objective of the experiment is to collect samples that come close to an everyday communicative situation. As it is well known, the more obstacles or atypicalities the participants in such research are faced with, the more unlikely the collected data will represent the language of everyday communication.

The problems that spontaneous speech pose to the analysis of prosodic phenomena are well known and widely discussed (see, for example, the collection of papers published by Sagisaka, Campbell & Higuchi 1997). Although there are techniques available to overcome some of these problems, as yet there exists no definite method that will prevent all of them from arising. The existence of such problems is, however, not an excuse for the preclusion of spontaneous speech in the study of prosody.

The need for information on the interaction between prosody and the organization of spontaneous speech is becoming more imperative, especially for the area of speech technology. It is known that segmental intelligibility, at present, does not constitute a problem in speech synthesis research any longer, due to the various techniques already available that produce highly intelligible synthetic speech. The major obstacle in this type of research is still the lack of comprehensive information on prosody. Inadequate knowledge of the prosodic system of a language will invariably result in the production of unnatural and even unpleasant synthetic speech. Information on how prosodic features are specifically employed to convey discourse structure can be applied to develop

algorithms to be used, for example, in the improvement of the comprehensibility of text generation systems.

Such information is also indispensable in the field of automatic recognition systems. The results in the present research could be applied, for example, to improve the capacity of understanding systems to recognize discourse structure more efficiently. One of the practical imports of such improvement is the development of better automatic summary producers that could interpret a more natural speech input.

From a more theoretical standpoint, the present study also contributes to the investigation of the formal characteristics of narrative structure in the broad field of narrative analysis. As it was mentioned in the introductory chapter, even though narrative structure is generally considered to be a defining quality of the narrative, no empirical and systematic evidence using spontaneous, non-elicited data has been presented so far to substantiate this assumption. It was demonstrated here that prosody can be considered a key element for the validation of this fundamental premise. This fact lends itself well to the conclusion that the consideration of the acoustic characteristics of discourse in the (more notional) field of narrative analysis is of irrefutable value. The present study can be regarded then as a contribution to the emergent line of research aimed at establishing interpretative links between linguistic phenomena and contextual processes, in a more interdisciplinary — rather than narrowly focused — approach.

The selection of Brazilian Portuguese for the purpose of the present analysis is not accidental. As mentioned in Chapter 1, despite the increasing empirical work on the prosody of discourse, this specific type of research in languages and cultures other than English and Dutch is still very much a *desideratum*. The present work is thus a

contribution to the (still incipient) study of the Brazilian Portuguese prosody, as well as to the research on storytelling in the Brazilian community, which is mostly centered on the non-acoustic aspects of the text. The importance of studies such as this rests on the fact that only with a large body of research on different aspects of prosody and (narrative) discourse carried out in different communities can one test the validity of the existing findings. This would consequently provide a deeper understanding of the various mechanisms of (narrative) discourse in different languages and cultures.

Throughout the analyses carried out in the present study, a number of findings requiring further investigation were reported. These findings, instead of reflecting consistent tendencies, constitute limited observations, leading predominantly to speculative conclusions. They are summarized below.

- Silent pauses that did not coincide with boundary pauses occurred sporadically in the data. The distribution of such pauses could be predicted on the basis of the linguistic elements that co-occurred with them: in almost 40% of the cases, pauses that did not occur between intonation units coincided with disruptive utterances (such as false starts, repeats and filled pauses); the other 60% co-occurred with content words (26%), discourse markers (22%) and other elements (13%). The duration of such pauses was distributed as follows: longer pauses were most often associated with discourse markers (35%) and disruptive utterances (30%). Functionally, most of these pauses reflected a cognitive process, such as re-elaboration of a statement or idea, planning ahead and/or lexical choices. Although these results are in agreement with previous studies on the same topic, further analyses need to be done.

- Speakers tended to increase speech rate at the introductory and conclusive sections of the narratives, suggesting the existence of a possible "narrative frame," marked by the manipulation of speech rate. Rapid speech at the introduction of a narrative may signal to the audience that a long turn is about to begin and that absolute attention is desired. In a highly competitive conversation, this acceleration in speech rate may be further interpreted as a means of "grabbing" the conversational turn. The high speech rate that occurs at codas may be a result of the (commonly) non-relevant role that they have in narratives, in general. As already discussed, fast speech is generally associated with information that is considered trivial. It would be interesting to investigate whether such a rate manipulation is a characteristic of any type of monological discourse and whether it has any effect on the perceptual level.
- A gradual lowering of pitch range at the end of the narratives was observed in some cases. This was interpreted as a possible signal of the narrative approaching the end. In view of the low number of cases in which this pitch range fading-away occurred, further analysis is needed for validating such an assumption.
- High pitch reset is indeed a reliable device in the signaling of narrative section boundaries, as discussed above. A closer analysis of this specific feature demonstrated that sometimes there exists a local manipulation of pitch range in the intonation units surrounding a narrative section boundary so that a high reset occurs. For example, the break that occurs between two sections that are characterized by an expanded pitch range value can be signaled by means of a

high pitch reset through the manipulation of either the last intonation unit of the first section or the first intonation unit of the proceeding section.

- This manipulation of pitch range at the local level appears also to be true when narrative sections as a whole are taken into account. For instance, when some intonation units in a narrative section regularly characterized by an average expanded pitch range are uttered in a low pitch range, the pitch range values of the other intonation units may be deliberately increased so as to make the average pitch range of the entire section high, as it is supposed to be. This would suggest that a compensatory effect might be at work withat the local level, with the purpose of producing an effect at the global level. Once again, further analysis over a more comprehensive corpus is in order before making generalizations.

It would be valuable to test these observations more accurately in order to validate the additional textual functions of prosody they suggest.

At this point, it is important to point out that prosody is not the only cue that can be used in the assignment of narrative structure. The interplay of an array of other linguistic and non-linguistic cues — such as syntactic, pragmatic, gestural, and visual cues — is what causes the audience to infer the intended global structure of a given text. It would be interesting to investigate the interaction of prosody with some other contextualizing cues in the process of demarcating discourse structure, as for example, the relation between prosodic and gestural cues. Several studies have already proposed that a close connection between intonation and gesture do exist, and that this can be observed when acquired language disorders are taken into account. Patients that suffer from such disorders regularly have the intonation pattern of their speech unaffected. It is

interesting that when the intonation is affected, gestural problems may appear (Cruttenden 1997). A tendency for certain gestures to be accompanied by specific intonation patterns has been also described by Cruttenden (1977): rising tones are often accompanied by eyebrows lifting, inclination of the head forward, shoulders raising, lengthy eye contact, hands lifting and/or palms upwards, suggesting that rising tones involve an increase in tension, whereas falling tones involve a decrease in tension. Valuable insight is still to be gained by systematically investigating the relationship between different prosodic features (not only pitch) and gestures in the demarcation of discourse structure.

References

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh, Edinburgh University Press.

Abrams, K. & Bever, T. G. (1969). Syntactic structure modifies attention during speech perception and recognition. *Quarterly Journal of Experimental Psychology* 21: 280-290.

Ainsworth, W. A. (1986). Pitch change as a cue to syllabification. *Journal of Phonetics* 14: 257-264.

Altenberg, B. (1987). *Prosodic Patterns in Spoken English: Studies in the Correlation between Prosody and Grammar for Text-to-Speech Conversion*. Lund, Lund University Press.

Armstrong, L. E. & Ward, I. C. (1926). *Handbook of English Intonation*. Leipzig, G.B. Teubner.

Armstrong, L. E. & Ward, I. C. (1931). *A Handbook of English Intonation*. Cambridge, W. Heffer & Sons.

Atkinson, J. E. (1973). *Aspects of intonation in speech: Implications from an experimental study of fundamental frequency*, The University of Connecticut.

Auer, P. (1990). Rhythm in telephone closings. *Human Studies* 13: 361-392.

Avesani, C. & Vayra, M. (1988). Discorso, segmenti di discorso e un' ipotesi sull' intonazione. *Proceeding of the Convegno Internazionale "Sull'Interpunzione"*, Florence.

Ayers, G., *et al.* (1995). Modelling intonation in dialogue. *Proceeding of the ICPhS*, Stockholm.

Ayers, G. M. (1992). Discourse functions of pitch range in spontaneous and read speech. *Proceeding of the Linguistic Society of America Annual Meeting.*

Baart, J. (1987). *Focus, Syntax, and Accent Placement. Towards a rule system for the derivation of pitch accent patterns in Dutch as spoken by humans and machines*, Leiden University.

Bachenko, J. & Fitzpatrick, E. (1990). A computational grammar of discourse-neutral prosodic phrasing in English. *Computational Linguistics* 16(3): 155-170.

Baken, R. J. (1987). *Measurement of Speech and Voice*. Nedham Heights, MA, Allyn and Bacon.

Bakhtin, M. (1995). *Maxismo e a Filosofia da Linguagem*. São Paulo, HUCITEC.

Bal, M. (1985). *Narratology: Introduction to the Theory of Narrative*. Toronto, University of Toronto Press.

Balkenhol, H., Hollien, H. & Hollien, P. (1979). Rhythm and pause as a means of expression in the Japanese arts of storytelling with special regard to Rakugo. In H.

Hollien & P. Hollien. *Current Issues in the Phonetic Sciences. Proc. of the IPS-77 Congress*. Amsterdam, Benjamins. 9: 235-243.

Baltiner, A., Kompe, R., Kiebling, A., Nöth, E. & Niemann, H. (1994). Can you tell apart spontaneous and read speech if you just look at prosody? In A. J. R. Ayuso & J. M. L. Soler. *Speech Recognition and Coding: New Advances and Trends*. Berlin, Springer: 321-324.

Bamberg, M. (1997). Positioning between structure and performance. *Journal of Narrative and Life History* 7 (1-4): 335-342.

Bamberg, M. & Damrad-Frye, D. (1991). On the ability to provide evaluative comments: Further explorations of children's narrative competencies. *Journal of Child Language* 18: 689-710.

Barik, H. C. (1977). Cross-linguistic study of temporal characteristics of different speech materials. *Language and Speech* 20: 116-126.

Barthes, R. (1975). An introduction to the structural analysis of narrative. *New Literary History* 6: 237-230.

Barthes, R. (1977). Introduction to the structural analysis of narratives. *Image-Music-Text*. London, Fontana.

Baum, S. R., Pell, M. D., Leonard, C. L. & Gordon, J. K. (1997). The ability of right- and left-hemisphere damaged individuals to produce and interpret prosodic cues marking phrasal boundaries. *Language and Speech* 40(4): 313-330.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language* 30: 644-663.

Beattie, G. (1980). Encoding units in spontaneous speech: some implications for the dynamics of conversation. In H. W. Dechert & M. Raupach. *Temporal Variables in Speech*. The Hague, Mouton: 131-143.

Beattie, G. (1983). *Talk: An Analysis of Speech and Non-Verbal Behavior in Conversation*. Milton Keyes, Open University Press.

Beattie, G. (1984). "Are there cognitive rhythms in speech?" - a reply to Power, 1983. *Language and Speech* 27: 193-195.

Beckman, M. E. (1997). A typology of spontaneous speech. In Y. Sagisaka, N. Campbell & N. Higuchi. *Computing Prosody: Computational Models for Processing Spontaneous Speech*. New York, Springer: 7-26.

Beckman, M. E. & Ayers, G. (1994). *Guidelines for ToBI labelling. Manuscript and accompanying speech materials*, Ohio State University.

Beer, M. (1910). Die Abhängigkeit der Lesezeit von psychologischen un sprachlichen Faktoren. *Zeitschrift für Psychologie* 56: 264-298.

Bernstein, B. (1962). Linguistic codes, hesitation phenomena and intelligence. *Language and Speech* 5: 31-46.

Blaauw, E. (1995). *On the perceptual classification of spontaneous and read speech.* Research Institute for Language and Speech, Utrecht University.

Black, J. B. & Wlilensky, R. (1979). An evaluation of story grammars. *Cognitive Science* 3: 213-230.

Black, J. W. (1950). The effect of room characteristics upon vocal intensity and rate. *Journal of the Acoustical Society of America* 22: 174-176.

Black, J. W., Tosi, O., Singh, S. & Takefuta, Y. (1966). A Study of pauses in oral reading of one's native language an English. *Language and Speech* 9: 237-241.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 17: 97-110.

Boersma, P. (1997). *Praat: doing phonetics by computer.* http://www.fon.hum.uva.nl/praat/.

Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language* 37: 83-96.

Bolinger, D. L. (1958). A theory of pitch accent in English. *Word* 14: 109-149.

Bolinger, D. L. (1972). Accent is predictable (if you're a mindreader). *Language* 48: 633-644.

Bolinger, D. L. (1978). Intonation across languages. In C. A. Ferguson & E. A. Moravcsik. *Universals of Human Language.* Stanford, Stanford University Press.

Bolinger, D. L. (1989). *Intonation and its Uses: Melody in Grammar and Discourse.* London, Edward Arnold.

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech* 8: 148-158.

Boomer, D. S. & Dittman, A. T. (1962). Hesitation pauses and juncture pauses in speech. *Language and Speech* 5: 215-220.

Boyce, S. & Menn, L. (1979). Peaks vary, endpoints don't: Implications for linguistic theory. *Proceeding of the the 1979 Meeting of the Berkeley Linguistic Society,* Linguistics Department, University of California at Berkeley.

Brazil, D. (1997). *The Communicative Value of Intonation in English.* Cambridge, Cambridge University Press.

Brazil, D., Coulthard, M. & Johns, C. (1980). *Discourse Intonation and Language Teaching.* London, Longman.

Brémond, C. (1973). *Logique du récit.* Paris, Seuil.

Brewer, W. F. (1985). The story schema: Universal and culture-specific properties. In D. A. Olsen, N. Torrance & A. Hildyard. *Literacy, Language and Learning.* Cambridge, Cambridge University Press: 167-194.

Brigance, W. N. (1926). How fast do we talk? *Quarterly journal of speech education* 12: 337-342.

Brotherton, P. (1979). Speaking and not speaking; Process for translating ideas into speech. In A. Siegman & S. Feldestein. *Of Time and Speech.* Hillsdale, N. J., Lawrence Erlbaum: 178-209.

Brown, G. (1983). Prosodic structure and the given/new distinction. In A. Cutler & D. R. Ladd. *Prosody: Models and Measurements.* Berlin, Heidelberg, New York, Tokyo, Springer-Verlag: 67-78.

Brown, G., Currie, K. & Kenworthy, J. (1980). *Questions of Intonation.* London, Croom Helm.

Brown, P. & Levinson, S. (1978). Universals in language usage: politeness phenomena. In E. N. Goody. *Questions and politeness. Strategies in social interaction.* Cambridge, Cambridge University Press: 56-289.

Brubaker, R. S. (1972). Rate and pause characteristics of oral reading. *Journal of Psycholinguistic Research* 1(2): 1972.

Bruce, G. (1982). Textual aspects of prosody in Swedish. *Phonetica* 39: 274-287.

Bruce, G. (1984). Aspects of declination in Swedish. *Working Papers in Phonetics* 24: 51-64.

Bruce, G., *et al.* (1997). On the analysis of prosody in interaction. In Y. Sagisaka, N. Campbell & N. Higuchi. *Computing Prosody: Computational Models for Processing Spontaneous Speech.* New York, Springer: 43-59.

Bruneau, T. J. (1973). Communicative silences: forms and functions. *The Journal of Communication* 23: 17-46.

Bruner, J. (1986). *Actual minds possible words.* Cambridge, Mass, Harvard University Press.

Bryan, K. L. (1989). Language prosody and the right hemisphere. *Aphasiology* 3(4): 285-299.

Butcher, A. (1981). *Aspects of the Speech Pause: Phonetic Correlates and Communicative Functions.* Institut für Phonetik, Universität Kiel.

Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research* 4: 75-87.

Butterworth, B. (1980). Evidence from pause in speech. In B. Butterworth. Language Production: Speech and Talk. London, Academic Press. 1: 155-176.

Butterworth, B. & Goldman-Eisler, F. (1979). Recent studies on cognitive rhythm. In A. W. Siegman & S. Feldstein. Of Speech and Time. Hillsdale, New Jersey, Lawrence Erlbaum.

Caffi, C. & Janney, R. V. O. (1994). Toward a pragmatics of emotive communication. *Journal of Pragmatics* 22: 325-373.

Cahn, J. (1995). The effect of pitch accenting on pronoun referent resolution. *Proceeding of the 33rd annual meeting of the association for computational linguistics.*

Cantril, H. & Allport, G. W. (1935). *The Psychology of Radio*. New York, Harper and Brothers.

Carter, R. & Simpson, P. (1982). The sociolinguistic analysis of narrative. *Belfast Working Papers in Linguistics* 6: 123-152.

Castilho, A. T. & Preti, D., Eds. (1986). *A Linguagem Falada Culta na Cidade de São Paulo: Elocuções Formais*. São Paulo, T. A. Queiroz.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind* 11: 63-65.

Chafe, W. (1989). Some things that narratives tell us about the mind. In B. B. A. Pellegrini. *Narrative Thought and Narrative Language*. Hillsdale, Lawrence Erlbaum.

Chafe, W. L. (1974). Language and consciousness. *Language* 50: 111-133.

Chafe, W. L. (1976). Giveness, constrativeness, definiteness, subjects, topics, and point of view. In C. Li. *Subject and Topic*. New York, Academic Press: 25-55.

Chafe, W. L. (1980). The deployment of consciousness in the production of a narrative. In W. L. Chafe. *The Pear Stories*. Norwood, N. J., Ablex: 9-50.

Chafe, W. L., Ed. (1980). *The Pear Story*. Norwood, N. J., Ablex.

Chatman, S. (1988). The representation of text types. *Textual Practices* 2: 22-29.

Clark, H. & Haviland, S. (1977). Comprehension and the given-new contract. In R. Freedle. *Discourse Production and Comprehension*. Hillsdale, N.J., Erlbaum Associates: 1-40.

Clark, H. H. & Clark, E. V. (1977). *Psychology and Language*. New York, Harcourt, Brace, Javanovich, Inc.

Clemmer, E. J. (1975). *Linguistic and Cognitive Aspects of Speech Hesitations in Schizophrenics and Normals*. Saint Louis, Saint Louis University.

Clemmer, E. J., O'Connell, D. C. & Loui, W. (1979). Rhetorical pauses in oral reading. *Language and Speech* 22: 397-405.

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika* 37: 256-266.

Cohen, A., Collier, R. & 't Hart, J. (1982). Declination: construct or intrinsic feature of speech pitch? *Phonetica* 39: 254-273.

Cohen, R. (1984). A computational theory of the function of clue words in argument understanding. *Proceeding of the COLING84*, Stanford.

Collier, R. (1975). The role of intonation in speech perception. In A. Cohenaud & S. Nooteboom. *Structure and Process in Speech Perception*. Berlin, Springer-Verlag: 107-123.

Collier, R. (1993). On the communicative function of prosody: some experiments. *IPO Annual Progress Report* 28: 67-75.

Collier, R., Piyper, J. R. d. & Sanderman, A. (1993). Perceived prosodic boundaries and their phonetic correlates. *Proceeding of the ARPA Workshop on Human Language Technology*, Plainsboro, New Jersey, USA, Morgan Kaufman Publishers.

Contini, M. & Boé, L. J. (1975). Etude quantitative de l'intonation en Français. *Proceeding of the 8th International Congress of the Phonetic Sciences*, Leeds.

Cook-Gumperz, J. (1992). Gendered context. In P. Auer & A. d. Luzio. *The Contextualization of Language*. Amsterdam, Benjamins: 177-198.

Cooper, W. & Sorensen, J. M. (1981). *Fundamental Frequency in Sentence Production*. New York, Springer-Verlag.

Cooper, W. E. & Sorensen, J. M. (1977). Fundamental frequency contours at syntactic boundaries. *Journal of the Acoustical Society of America* 62: 682-692.

Cosmides, L. (1983). Invariances in the acoustic expression of emotion during speech. *Journal of Experimental Psychology: Human Perception and Performance* 9(6): 864-881.

Couper-Kuhlen, E. (1986). *An Introduction to English Prosody*. London, Edward Arnold.

Couper-Kuhlen, E. (1992). Contextualizing discourse: the prosody of interactive repair. In P. Auer & A. d. Luzio. *The Contextualization of Language*. Amsterdam, Benjamins: 337-364.

Couper-Kuhlen, E. (1996). Intonation and clause combining in discourse: the case of *because*. *Pragmatics* 6(3): 389-426.

Couper-Kuhlen, E. (1996). The prosody of repetition: on quoting and mimicry. In E. Couper-Kuhlen & M. Selting. *Prosody in Conversation*. Cambridge, Cambridge University Press.

Couper-Kuhlen, E. (1998). Coherent voicing. On prosody in conversational reported speech. *Working Papers in Interaction and Linguistic Structures* 1: 1-26.

Couper-Kuhlen, E. & Selting, M. (1996). Towards an interactional perspective on prosody and prosodic perspective on interaction. In E. Couper-Kuhlen & M. Selting. *Prosody in Conversation*. Cambridge, Cambridge University Press: 11-56.

Couthard, M. (1992). The significance of intonation in discourse. In Coulthard-Malcolm. *Advances in Spoken Discourse Analysis*. London, Routledge: 35-49.

Cruttenden, A. (1997). *Intonation*. Cambridge, Cambridge University Press.

Crystal, D. (1969). *Prosodic Systems and Intonation in English*. London, Cambridge University Press.

Crystal, D. (1975). *The English Tone of Voice*. London, Arnold.

Crystal, T. H. & House, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America* 88: 101-112.

Culicover, P. W. & Rochemont, M. (1983). Stress and focus in English. *Language* 59(1): 123-165.

Currie, K. L. (1980). An initial search for tonics. *Language and Speech* 23: 329-350.

Cutler, A. (1983). Speaker's conception of the function of prosody. In A. Cutler & D. R. Ladd. *Prosody: Models and Measurements*. Berlin, Heidelberg, New York, Tokyo, Springer-Verlag: 79-92.

Cutler, A. (1997). Prosody and the structure of the message. In Y. Sagisaka, N. Campbell & N. Higuchi. *Computing Prosody: Computational Models for Processing Spontaneous Speech*. New York, Springer: 61-65.

Cutler, A. & Clifton, C. E. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis. *Attention and Performance X: Control of Language Processes*. Hillsdale, NJ, Erlbaum: 183-196.

Cutler, A., Dahan, D. & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech* 40(2): 141-201.

Cutler, A. & Isard, S. D. (1980). The production of prosody. In B. Butterworth. *The Production of Prosody*. London, Academic Press. 1: 245-269.

Cutler, A. & Ladd, D. R., Eds. (1983). *Prosody: Models and Measurements*. Springer Series in Language and Communiation. Berlin, Heidelberg, New York, Tokyo, Springer-Verlag.

Cutler, A. & Pearson, M. (1986). On the analysis of prosodic turn-taking cues. In C. Johns-Lewis. *Intonation in Discourse*. London, Croom Helm: 139-155.

Dalton, P. & Hardcastle, W. J. (1977). *Disorders of Fluency an Their Effects on Communication*. London.

Darnell, R. (1974). Correlates of Cree narrative performance. In R. Bauman & J. Sherzer. *Explorations in the ethnography of speaking*. Cambridge, Cambridge University Press: 315-336.

de Johnson, T. H., O'Connell, D. C. & Sabin, E. J. (1979). Temporal analyses of English and Spanish narratives. *Bulletin of the Psychonomic Society* 13: 347-350.

Dechert, H. W. & Raupach, M. (1980). *Temporal variables in speech*. New York, The Haugue.

Deese, J. (1984). *Thought into Speech: The Psychology of a Language*. Englewood Cliffs, NJ, Prentice-Hall.

di Cristo, A. & Hirst, D. J. (1986). Modelling French micromelody: analysis and synthesis. *Phonetica* 43: 11-30.

Dolbec, J. & Demers, M. (1993). Prosodie et fonction discursive: le cas du là en Français Québécois. *Proceeding of the XVth International Congress of Linguistics*, Saint-Foy, Le Press de L'Université Laval.

Douglas-Cowie, E. & Cowie, R. (1998). Intonational Settings as Markers of Discourse Units in Telephone Conversations. *Language and Speech* 41 (3-4): 351-374.

Dryer, M. S. (1996). Focus, pragmatic pressuposition and activated propositions. *Journal of Pragmatics* 26: 475-523.

Duez, D. (1982). Silent and non-silent pauses in three speech styles. *Language and Speech* 25: 11-28.

Eady, S. J. & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America*(80): 402-415.

Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R. & Lotts, D. W. (1986). Acoustical characteristics of setential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech* 29: 233-251.

El-Menoufy, A. (1988). Intonation and Meaning in Spontaneous Discourse. In J. Benson, M. J. Cummings & G. S. Wlliam. *Linguistics in a Systemic Perspective*. Amsterdam, Benjamins: 1-26.

Enkevist, N. E. & Nordström, H. (1978). On textual aspects of intonation in Finland-Swedish newscasts. *Studia Linguistica* 32: 63-79.

Erbaugh, M. S. (1996). A uniform pause and error strategy for native and non-native speakers. *Proceeding of the International Conference on Spoken Language Processing*, Philadelphia.

Essen, O. v. (1949). Sprech tempo als Ausdruck psychischen Geschehens. *Zeitschrift für Phonetik* 3: 317-341.

Ewan, W. (1979). Can intrinsic vowel fo be explained by source/tract coupling? *Journal of the Acoustical Society of America* 66: 358-362.

Fairbanks, G. & Hoaglin, L. W. (1940). An experimental study of the durational characteristics of the voice during the expression of emotion. *Speech Monographs* 7: 85-90.

Faure, G., Hirst, D. J. & Chafcouloff, M. (1980). Rhythm in English: Isochronism, pitch, and perceived stress. In L. R. Waugh & C. H. v. Schooneveld. *The Melody of Language*. Baltimore, University Park Press: 71-79.

Fernandes, N. (1976). *Contribuição para uma Ánalise Instrumental da Acentuação e Intonação do Português*, Universidade de São Paulo.

Fisher, W. (1987). *Human communication as narration: toward a philosophy of reason, value, and action*. Columbia, University of South Carolina Press.

Fleischman, S. (1990). *Tense and Narrativity: From Medieval Performance to Modern Fiction*. Austin, University of Texas Press.

Fon, J. (1999). Speech rate as a reflection of variance and invariance in conceptual planning in storytelling. *Proceeding of the ICPhS*.

Fónagy, I. (1978). A new method of investigating the perception of prosodic features. *Language and Speech* 21: 34-49.

Fónagy, I. & Magdics, K. (1960). Speed of utterance in phrases of different lenghts. *Language and Speech* 4: 179-192.

Ford, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan. *The Mental Representation of Grammatical Relations*. Cambridge, Mass., MIT Press.

Fowler, C. A. & Housum, J. (1987). Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and used of the distinction. *Journal of Memory and Language* 26: 489-504.

Fox Tree, J. E. (1993). *Comprehension After Speech Disfluencies*. Stanford, CA, Stanford University.

Fröschels, E. (1920). Untersuchungen über das Sprechtempo. *Monatsschrift für Ohrenheilkunde und Laryngo-Rhinologie* 54: 867-871.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech* 1: 126-152.

Fry, D. B. (1968). Prosodic phenomena. In B. Malmberg. *Manual of Phonetics*. Amsterdam, North-Holland Publishing Company: 265-410.

Fuchs, A. (1984). 'Deaccenting' and 'default accent'. In D. Gibbon & H. Richter. *Intonation, Accent and Rhythm*. New York, De Gruyter: 134-164.

Fujisaki, H. (1997). Prosody, models, and spontaneous speech. In Y. Sagisaka, N. Campbell & N. Higuchi. *Computing Prosody: Computational Models for Processing Spontaneous Speech*. New York, Springer: 27-42.

Fujisaki, M., Hirosi, K. & Ohta, K. (1979). Acoustic features of the fundamental frequency of connected speech in Japanese. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 13: 163-172.

Gale, W., Church, K. W. & Yarowsky, D. (1992). Estimating upper and lower bounds on the performance of word-sense disambiguation programs. *Proceeding of the 30th Association for Computational Linguistics*, Newark, Delaware.

Garding, E. (1967). Prosodiska drag i spontant och uppläst tal. In G. Holm. *Svenkst talsprak*. Stockholm, Almqwist & Wiksell: 40-85.

Garding, E. (1982). Prosodic expressions and pragmatic categories. *Working Papers of Lund University Department of Linguistics* 22: 117-135.

Garnica, O. K. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson. *Talking to Children: Language Input and Acquisition*. Cambridge, Cambridge University Press. 63-88.

Gee, J. P. (1985). The narrative of experience in the oral style. *Journal of Education* 167: 9-35.

Gee, J. P. (1986). Units in the production of narrative discourse. *Discourse Processes* 9: 391-422.

Gee, J. P. (1989). Literariness, formalism and sense making: the line and stanza of human thought. *Journal of Education* 171: 61-74.

Gee, J. P. & Grosjean, F. (1983). Performance structure: A psycholinguistic and linguistic apprasial. *Cognitive Psychology* 15: 411-458.

Gee, J. P. & Grosjean, F. (1984). Empirical evidence for narrative structure. *Cognitive Science* 8: 59-85.

Gee, J. P. & Kegl, J. A. (1983). Narrative/story structure, pausing and ASL. *Discourse Processes* 9: 391-422.

Gelfer, C. E., Harris, K., Collier, R. & Baer, T. (1983). Speculations on the control of fundamental frequency declination. *Status Report on Speech Research* 76: 51-68.

Geluykens, R. & Swerts, M. (1993). Local and global prosodic cues to discourse organization in dialogues. *Proceeding of the ESCA Workshop on Prosody*, Lund, Sweden, Dept. of Linguistics and Phonetics, Lund University.

Geluykens, R. & Swerts, M. (1994). Prosodic cues to discourse boundaries in experimental dialogues. *Speech Communication* 15: 69-77.

Genette, G. (1972). *Figures III*. Paris, Seuil.

Genette, G. (1980). *Narrative Discourse*. Ithaca, Cornell University Press.

Gibbon, D. (1987). Intonation and discourse. In J. S. Petöfi. *Text and Discourse Constituition*. Berlin, de Gruyter: 3-25.

Gimson, A. C. (1956). The linguistic relevance of stress in English. *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft* 9: 113-149.

Givón, T. (1982). Tense-aspect-modality: The creole prototype and beyond. In P. Hopper. *Tense-Aspect: Between Semantics and Pragmatics*. Amsterdam, John Benjamins.

Glenn, C. G. (1978). The role of episodic structure and of story length on children's recall of single stories. *Journal of Verbal Learning and Verbal Behavior* 17: 229-247.

Glukhov, A. A. (1975). Statistical analysis of speech pauses for Romance and Germanic languages. *Soviet Physics. Acoustics* 21: 71-72.

Goffman, E. (1955). On face-work: An analysis of ritual elements in social interaction. *Psychiatry* 18: 213-231.

Goldman-Eisler, F. (1956). The determinants of the rate of speech output and their mutual relations. *Journal of the Psychosomatic Research* 1: 137-143.

Goldman-Eisler, F. (1958). The predictability of words in contexts and the length of pauses in speech. *Language and Speech* 1: 226-231.

Goldman-Eisler, F. (1958). Speech analysis and mental processes. *Language and Speech* 1: 59-75.

Goldman-Eisler, F. (1961). A comparative study of two hesitation phenomena. *Language and Speech* 4: 18-26.

Goldman-Eisler, F. (1961). Continuity of speech utterance, its determinants and its significance. *Language and Speech* 4: 220-231.

Goldman-Eisler, F. (1961). The distribution of pause duration in speech. *Language and Speech* 4: 232-237.

Goldman-Eisler, F. (1961). The rate of changes in the rate of articulation. *Language and Speech* 4: 171-174.

Goldman-Eisler, F. (1967). Sequential temporal patterns and cognitive processes in speech. *Language and Speech* 1967: 122-132.

Goldman-Eisler, F. (1968). *Psycholinguistics: experiments in spontaneous speech.* London, New York, Academic Press.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech* 15: 103-113.

Goodwin, M. H. (1996). Informing and announcements in their environment: prosody within a multi-activity. In E. Couper-Kuhlen & M. Selting. *Prosody in Conversation.* Cambridge, Cambridge University Press: 436-461.

Graddol, D. (1986). Discourse specific pitch behavior. In C. Johns-Lewis. *Intonation in Discourse.* London, Croom Helm, College-Hill: 221-272.

Graddol, D. & Swann, J. (1989). *Gender Voices.* Oxford, Basil Blackwell.

Green, D. W. (1977). The immediate processing of sentences. *Quarterly Journal of Experimental Psychology* 29: 135-146.

Greimas, A. J. (1970). *Du Sens.* Paris, Seuil.

Grosjean, F. (1980). Temporal variables within and between languages. In H. Dechert & M. Raupach. *Towards a Cross-Linguistic Assessment of Speech Production.* Frankfurt, Lang.

Grosjean, F. (1983). How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics* 21: 501-529.

Grosjean, F. & Collins, M. (1979). Breathing, Pausing and Reading. *Phonetica* 36: 98-114.

Grosjean, F. & Deschamps, A. (1972). Analyse des variables temporelles du français spontané. *Phonetica* 26: 126-156.

Grosjean, F. & Deschamps, A. (1973). 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* 28: 191-226.

Grosjean, F. & Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica* 14: 141-148.

Grosz, B. & Hirschberg, J. (1992). Some intonational characteristics of discourse structure. *Proceeding of the International Conference on Spoken Language Processing*, Banff.

Grosz, B. & Sidner, C. L. (1986). Attentions, intentions, and the structure of discourse. *Computational Linguistic* 85: 363-394.

Gülich, H. & Quasthoff, U. (1985). Narrative analysis. In T. v. Djik. *Handbook of Discourse Analysis.* London, Academic Press. 2: 169-197.

Gülich, H. & Quasthoff, U. (1986). Story-telling in conversation: Cognitive and interactive aspects. *Poetics* 5: 217-241.

Günthner, S. (1997). The contextualization of affect in reported dialogues. In S. Niemier & R. Dirven. *The Language of Emotions*. Amsterdam, John Benjamins: 247-275.

Gussenhoven, C. (1983). Focus, mode and nucleus. *Journal of Linguistics* 19: 377-417.

Gussenhoven, C. & Rietveld, A. C. M. (1988). Fundamental frequency declination in Dutch: testing three hypothesis. *Journal of Phonetics* 16: 355-369.

Hakoda, K. & Sato, H. (1980). *Prosodic rules in connected speech synthesis*, Translation of the Institute of Electronics and Communication Engineers 63-D.

Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. The Hague, Mouton.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English. Part 2. *Journal of Linguistics* 2: 199-244.

Halliday, M. A. K. (1970). *A Course in Spoken English: Intonation*. London, Oxford University Press.

Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London, Edward Arnold.

Halliday, M. A. K. & Hassan, R. (1976). *Cohesion in English*. London, Longman.

Hanley, T. D. & Snidecor, J. C. (1967). Some acoustic similarities among languages. *Phonetica* 17: 141-148.

Hanley, T. D., Snidecor, J. C. & Ringel, R. L. (1966). Some acoustic differences among languages. *Phonetica* 14: 97-107.

Hansson, P. (1998a). Pausing in Spontaneous Speech. *Proceeding of the FONETIK 98, The Swedish Phonetics Conference*, Stockholm University.

Hansson, P. (1998b). Pausing in Swedish Spontaneous Speech. *Proceeding of the Seventh Australian International Conference on Speech Science and Technology*, Sydney.

Hansson, P. (1999). Discourse markers in dialogue. *Proceeding of the FONETIK 99: the 12th Swedish Phonetics Conference*, Göteborg University.

Hargreaves, W. A. & Starkweather, J. A. (1959). Collection of temporal data with duration tabulator. *J. Exp. Anal. Behav.* 2: 179.

Haviland, S. E. & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior* 13: 512-520.

Hawkins, P. R. (1971). The syntactic location of hesitation pauses. *Speech and Language* 14: 277-288.

Hearst, M. A. (1993). *TextTiling: A quantitative approach to discourse segmentation*. Berkeley, University of California.

Hearst, M. A. (1994). Multi-paragraph segmentation of expository text. *Proceeding of the 32nd ACL*.

Heath, S. B. (1982). What no bedtime story means: narrative skills at home and school. *Language in Society* 11: 49-76.

Heath, S. B. (1983). *Ways with Words: Language, Life and Work in Communities and Classrooms*. Cambridge, Cambridge University Press.

Hegedüs, L. (1957). Sprechtempoanalysen im Ungarischen. *Zeitschift für Phonetik* 10: 8-20.

Heinitz, W. (1921). Die Bewertung der Dauer in phonetischen Aufnahmen. *Vox* 153.

Henderson, A., Goldman-Eisler, F. & Skarbeck, A. (1965). Temporal patterns of cognitive activity and breath control in speech. *Language and Speech* 8: 236-242.

Henderson, A., Goldman-Eisler, F. & Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech. *Language and Speech* 9: 207-216.

Henze, R. (1953). Experimentelle Untersuchungen zur Phänomenologie der Sprechgeschwindigkeit. *Zeitschrift für experimentelle und angewandte Psychologie* 1: 214-243.

Hermes, D. J. (1988). Measurement of pitch by subharmonic summation. *J. Acout. Soc. Am.* 83(1): 257-264.

Hess, W. (1983). *Pitch Determination of Speech Signals*. Berlin, Spiger-Verlag.

Hicks, D. (1990). Narrative skills and genre knowledge: ways of telling in the primary school grades. *Applied Psycholinguistics* 11: 83-104.

Hicks, D. (1991). Kinds of narrative: genre skills among first graders from two communities. In A. McCabe & C. Peterson. *Developing Narrative Structure*. Hillsdale, NJ, Erlbaum: 55-87.

Hieke, A. E., Kowal, S. & O'Connell, D. C. (1983). The trouble with "articulatory pauses". *Language and Speech* 26: 203-214.

Hillenbrand, J., Getty, L. A., Clarck, M. J. & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* 97(5): 3099-3111.

Hinds, J. (1979). Organizational patterns in discourse. In T. Givon. *Discourse and Syntax*. New York, Academic Press. 12: 135-157.

Hirschberg, J. & Grosz, B. (1992). Intonation features of local and global discourse structure. *Proceeding of the DARPA Workshop on Spoken Language Systems*, Arden House.

Hirschberg, J. & Litman, D. (1993). Empirical studies on the disambiguation of cue phrases. *Computational Linguistics* 19.

Hirschberg, J., Nakatani, C. H. & Grosz, B. J. (1995). Conveying discourse structure through intonation variation. *Proceeding of the ESCA Workshop on Spoken Dialogue Systems: Theories and Applications*, Visgo, Denmark, ESCA.

Hirschberg, J. & Pierrehumbert, J. (1986). The intonational structuring of discourse. *Proceeding of the Twenty-fourth Annual Meeting of the Association for Computational Linguistics*.

Hirst, D. & di Cristo, A., Eds. (1998). *Intonation systems : a survey of twenty languages.* New York, Cambridge University Press.

Hopper, P. (1979). Aspect and foregrounding in discourse. In T. Givón. *Discourse and Syntax.* New York, Academic Press. 12.

Hopper, P. & Thompson, S. (1956). Transitivity in language and discourse. *Language* 56: 251-299.

Horne, M. (1991). Phonetic correlates of the 'given/new' parameter. *Proceeding of the International Congress of Phonetic Sciences 91*, Aix-en-Provence.

Horne, M., Bruce, G., Hansson, P. & Filipsson, M. (1998). Boundary-signaling in discourse: the case of Swedish *men*, an ambiguous cue word. *Proceeding of the FONETIK 98, The Swedish Phonetics Conference*, Stockholm.

Horne, M., Hansson, P., Bruce, G., Frid, J. & Filipsson, M. (1999). Discourse markers and the segmentation of spontaneous speech: The case of Swedish *men* 'but/and/so'. *Working Papers of the Department of Linguistics and Phonetics, Lund University* 47: 123-140.

Howell, P. & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication* 10: 163-169.

Hudson, J. A. & Shapiro, R. (1991). From knowing to telling: The development of children's scripts, stories, and personal narratives. In A. McCabe & C. Peterson. *Developing Narrative Structure.* Hillsdale, NJ, Erlbaum: 89-136.

Huss, V. (1978). English word stress in post-nuclear position. *Phonetica* 35: 86-105.

Hwang, C. H. & Schubert, L. K. (1992). Tense trees as the 'fine structure' of discourse. *Proceeding of the the 30th Annual Meeting of the ACL.*

Hyman, L. M. (1975). *Phonology. Theory and Analysis.* New York, Holt, Rinehart and Winston.

Hymes, D. (1981). "In vain I tried to tell you": Essays of Native American ethnopoetics. Philadelphia, University of Pennsylvania Press.

Jackendorff, R. S. (1972). *Sentence Interpretation in Generative Grammar.* Cambridge MA, MIT Press.

Jassem, W. (1971). On the pitch and compass of the speaking voice. *Journal of the International Phonetic Association* 1: 59-68.

Jassem, W. & Demenko, G. (1996). On Extracting Linguistic Information from Fo Traces. In C. Johns-Lewis. *Intonation in Discourse.* London, Croom Helm, College-Hill: 1-17.

Jefferson, G. (1978). Sequential aspects of storytelling in conversation. In J. Schenkein. *Studies in the organization of conversational interaction.* New York, Academic Press: 219-248.

Jespersen, O. (1913). *Lehrbuch der Phonetik.* Leipzig.

Johns-Lewis, C. M. (1986). Prosodic differentiation of discourse modes. In C. M. Johns-Lewis. *Intonation in Discourse*. London, Croom Helm, College-Hill: 199-219.

Johnson, K. (1997). *Acoustic and Auditory Phonetics*. Cambridge, Blackwell.

Johnson, T. H. d., O'Connell, D. C. & Sabin, E. J. (1979). Temporal analysis of English and Spanish narratives. *Bulletin of the Psychonomic Society* 13: 347-350.

Johnstone, B. (1990). *Stories, Community and Place: Narratives from Middle America*. Bloomington, Indiana University Press.

Johnstone, B. (1996). *The Linguistic Individual*. New York, Oxford, Oxford University Press.

Jones, D. (1909). *Intonation Curves*. Leipzig, G.B. Teuber.

Kallmeyer, W. & Schütze, F. (1977). Zur Konstotution von Kommunikationsschemata der Sachverhaltsdarstellung. In D. Wegner. *Gesprächsanalysen*. Hamburg, Buske: 159-274.

Kernan, K. (1977). Semantic and expressive elaboration in children's narratives. In S. Ervin-Tripp & C. Mitchell-Kernan. *Child Discourse*. New York, Academic Press: 91-102.

Kingdon, R. (1958). *The Groundwork of English Intonation*. London, Longman, Green and Company.

Kintsch, W. (1974). *The Representation of Meaning in Memory*. Hillsdale, NJ, Lawrence Erlbaum Associates.

Kintsch, W. & Greene, E. (1978). The role of culture-specific schemata in the comprehension and recall of stories. *Discourse Processes* 1: 1-13.

Kirschenblatt-Gimblett, B. (1974). The concept and varieties of narrative performance in east european Jewish culture. In R. Bauman & S. Sherzer. *Explorations in the ethnography of speaking*. Cambridge, Cambridge University Press: 283-308.

Koopmans-van Beinum, F. J. (1990). Spectro-temporal reduction and expansion in spontaneous and read speech: The role of focus words. *Proceeding of the International Conference on Spoken Language Processing*.

Koopmans-van Beinum, F. J. (1991). Spectro-temporal reduction and expansion in spontaneous speech and read text: Focus words versus non-focus words. *Proceeding of the ESCA Workshop on Phonetics and Phonology of Speaking Styles*.

Koopmans-van Beinum, F. J. & Van Donzel, M. E. (1996). Discourse structure and its influence on local speech rate. *Proceeding of the International Conference on Spoken Language Processing*, Philadelphia.

Kowal, S. & O'Connell, D. C. (1987). Some temporal aspects of stories told while or after watching a film. *Bulletin of the Psychonomic Society* 25(5): 364-366.

Kowal, S., O'Connell, D. C. & Sabin, E. J. (1975). Development of temporal patterning and vocal hesitation in spontaneous narratives. *Journal of Psycholinguistic Research* 4: 195-207.

Kowal, S., Wiese, S. & O'Connell, D. (1983). The use of time in storytelling. *Language and Speech* 26(4): 377-392.

Kowal, S. H., O'Connell, D. C., O'Brien, E. A. & Bryant, E. T. (1975). Temporal aspects of reading and speaking: Three experiments. *American Journal of Psychology* 88: 549-569.

Krause, M. (1984). Recent developments in speech signal pitch extraction. In D. Gibbon & H. Richter. *Intonation, Accent and Rhythm: Studies on Discourse Phonology.* Berlin, de Gruyter: 243-252.

Kreiman, J. (1982). Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics* 10: 163-175.

Kumpf, L. E. (1987). The use of pitch phenomena in the structuring of stories. In R. Tomlin. *Coherence and Grounding in Discourse.* Amsterdam, John Benjamins: 189-216.

Kuno, S. (1972). Functional sentence perspective. *Linguistic Inquiry* 3: 269-320.

Laan, G. P. M. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a reading speaking style. *Speech Communication* 22: 43-65.

Labov, W. (1971). The study of language in its social context. In J. Fishman. *Advances in the Sociology of Language.* The Hague, Mouton.

Labov, W. (1972). The transformation of experience in narrative syntax. *Language in the inner City.* Philadelphia, University of Pennsylvania Press: 354-98.

Labov, W. & Waletzky, J. (1967). Narrative analysis: oral versions of personal experience. In J. Helms. *Essays on the Verbal and Visual Arts.* Seattle, University of Washington Press.

Ladd, D. R. (1980). *The Structure of Intonational Meaning, Evidence from Speech.* Bloomington/London, Indiana University Press.

Ladd, D. R. (1984). Declination: a review and some hypotheses. *Phonology Yearbook* 1: 53-74.

Ladd, D. R. (1988). Declination "reset" and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84: 530-544.

Ladd, D. R. (1996). *Intonational Phonology.* Cambridge, Cambridge University Press.

Ladd, D. R. & Cutler, A. (1983). Models and measurements in the study of prosody. In A. Cutler & D. R. Ladd. *Prosody: Models and Measurements.* Berlin, Heidelberg, New York, Tokyo, Springer-Verlag: 1-10.

Ladd, D. R. & Silverman, K. E. A. (1984). Vowel intrinsic pitch in connected speech. *Phonetica* 41: 31-40.

Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus, and the mental representation of discourse referents.* Cambridge, Cambridge University Press.

Lass, N. J. & Deem, J. F. (1972). Temporal patterns of speech rate alternations. *Proceeding of the 7th International Congress of Phonetic Sciences*, Den Haag.

Lavoie, J. (1994). La fréquence intrinsèque des voyelles en français québécois. *Proceeding of the 8e Journées de linguistique.*

Lay, C. H. & Paivio, A. (1969). The effects of task difficulty and anxiety on hesitation in speech. *Canadian Journal of Behavioral Science* 1: 25-37.

Lea, W. A. (1980). Prosodic aids to speech recognition. In W. A. Lea. *Trends in Speech Recognition*. Englewood Cliffs, Prentice-Hall: 166-205.

Leech, G. & Short, M. (1981). *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. London, Longman.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MIT.

Lehiste, I. (1975). The phonetic structure of paragraphs. In A. Cohen & S. Nooteboom. *Structure and Process in Speech Perception*. Berlin, Springer-Verlag: 195-206.

Lehiste, I. (1976). Suprasegmental features of speech. In J. Lass. *Contemporary Issues in Experimental Phonetics*. New York, Academic Press: 225-239.

Lehiste, I. (1979). Perception of sentence and paragraph boundaries. In B. Lindblom & S. Öhman. *Frontiers of Speech Communication Research*. London, Academic Press: 191-201.

Lehiste, I. (1982). Some phonetic characteristics of discourse. *Studia Linguistica* 36(2): 117-130.

Lehiste, I. (1996). Suprasegmental features of speech. In R. D. Ladd. *Principles of Experimental Phonetics*. St. Louis, Mosby: 226-244.

Lehiste, I. & Peterson, G. E. (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America* 33(4): 419-425.

Lehiste, I. & Wang, W. S.-Y. (1977). Perception of sentence and paragraph boundaries with and without semantic information. In W. U. Dressler & O. E. Pfeiffer. *Phonologica 1976*. Innsbruck, Institut für Sprachwissenschaft der Universität Innsbruck: 277-283.

Lehnert, W. G. (1981). Plot unit and narrative summarizations. *Cognitive Science* 4: 293-331.

Lerner, G. H. (1996). On the place of linguistic resources in the organization of talk-in-interaction 'second person' reference in multi-party conversation. *Pragmatics* 6(3): 281-294.

Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge, Cambridge University Press.

Levin, H., Schaffer, C. & Snow, C. (1982). The Prosodic and Paralinguistic Features of Reading and Telling Stories. *Language and Speech* 25(1): 43-54.

Levin, H. & Silverman, I. (1965). Hesitation phenomena in children's speech. *Language and Speech* 8: 67-85.

Levin, H., Silverman, I. & Ford, B. (1965). Hesitations in children's speech during explanation and description. *Journal of Verbal Learning and Verbal Behavior* 6: 560-564.

Levin, H., Silverman, I. & Ford, B. L. (1967). Hesitation in children's speech during explanation and description. *Journal of Verbal Learning and Verbal Behavior* 6: 560-564.

Lewis, D. & Tiffin, J. (1933). A psychophysical analysis of individual differences in speech. *Archives of Speech* 1: 43-60.

Liberman, M. & Sag, I. (1974). Prosodic form and discourse function. *Proceeding of the Papers from the tenth regional meeting*, Chicago Linguistic Society.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America* 32: 451-454.

Lieberman, P. (1967). *Intonation, Perception and Language*. Cambridge, MA, MIT Press.

Lieberman, P. (1974). A study of prosodic features. In T. A. Sebeok. *Current Trends in Linguistics*. The Hague, Mouton. 12: 2419-2449.

Lieberman, P. & Tseng, C. Y. (1981). On the fall of the declination theory: breath group versus declination as the base form for intonation. *Journal of the Acoustical Society of America* Supl. 1 67(S 63).

Linde, C. (1979). Focus of attention and the choice of pronouns in discourse. In T. Givón. *Discourse and Syntax*. New York, Academic Press: 337-354.

Litman, D. J. & Passonneau, R. J. (1993). Empirical evidence for intention-based discourse segmentation. *Proceeding of the ACL Workshop on Intentionality and Structure in Discourse Relations*.

Litman, D. J. & Passonneau, R. J. (1995). Combining Multiple Knowledge Sources for Discourse Segmentation. *Proceeding of the Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics (ACL-95)*, Cambridge, MA.

Local, J. (1992). Continuing and restarting. In P. Auer & A. d. Luzio. *The Contextualization of Language*. Amsterdam, Benjamins: 273-296.

Lote, G. (1911). Le silence et la pronunciation dans l'alexandrin français. *Revue de phonétique* 1: 114-137, 233-259.

Lounsbury, F. G. (1954). Transitional probability, linguistic structure and systems of habit-family hierarchies. In C. E. O. T. A. Sebok. *Psycholinguistics*. Baltimore.

Lumley, F. H. (1933). Rates of speech in radio speaking. *Quarterly Journal of Speech* 19: 393-403.

Lyons, J. (1972). Phonemic and non-phonemic phonology. *International Journal of American Linguistics* 28(2): 127-134.

Maclay, H. & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word* 15: 19.

Maclean, M. (1988). *Narrative as Performance: A Baudelairean Experiment.* London, Routledge.

Maeda, S. (1976). *A Characterization of American English Intonation.* Massachusetts, MIT.

Magno-Caldognetto, E., Ferrero, F. E., Lavagnoli, C. & Vagges, K. (1978). Fo contours of statements, yes-no questions and wh-questions of two regional varieties of Italian. *Journal of Italian Linguistics* 3: 57-68.

Major, R. (1981). Stress-timing in Brazilian Portuguese. *Journal of Phonetics* 9(3): 343-351.

Major, R. (1985). Stress and rhythm in Brazilian Portuguese. *Language* 61(2): 259-282.

Malinowski, B. (1923). Phatic communion. In J. Laver & F. Hutcheson. *Communication in face-to-face-interaction.* Harmondsworth, Peguin: 146-152.

Mandler, J. M. & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology* 9: 111-151.

Marchuschi, L. A. (1987). *Marcadores Conversacionais no Português Brasileiro: Formas, Posições e Funções.* Recife/Freiburg, Unpublished paper.

Marcus, M. P. & Hindle, D. (1985). A computational account of extra categorical elements in Japanese. In S. Y. Kuroda. *Papers presented at the First SDF Workshop in Japanese Syntax.* La Jolla, System Development Foundation.

Martin, J. G. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior* 9: 75-78.

Massini, G. A. (1991). *A Duração no Estudo do Acento e do Ritmo do Português.* Campinas, Unicamp.

Massini-Cagliari, G. (1992). *Acento e Ritmo.* São Paulo, Contexto.

Meinhold, G. (1972). Allgemeine Probleme der Sprechgeschwindigkeit. *Zeitschrift für Phonetik* 25: 492-505.

Menn, L. & Boyce, S. (1982). Fundamental frequency and discourse structure. *Language and Speech* 25(4): 341-379.

Michaels, S. (1981). Sharing time: children's narrative styles and differential access to literacy. *Language in Society* 10: 423-442.

Michaels, S. (1991). The dismantling of narrative. In A. McCabe & C. Peterson. *Developing Narrative Structure.* Hillsdale, NJ, Erlbaum: 303-351.

Michaels, S. & Collins, J. (1984). Oral discourse styles: classroom interaction and the acquisition of literacy. In D. Tannen. *Coherence in Spoken and Written Discourse.* Norwood, NJ, Ablex: 219-244.

Miller, G. A., Galanter, E. & Pribram, K. H. (1960). *Plans and the Structure of Behavior.* New York, Holt, Rinehart and Wiston Inc.

Moraes, J. (1998). Intonation in Brazilian Portuguese. In D. Hirst & A. di Cristo. *Intonation systems : a survey of twenty languages*. New York, Cambridge University Press.

Moraes, J. & Espesser, R. (1988). La hiérarchie des indices prosodiques de l'accent en portugais. *Proceeding of the the First International Conference on Experimental Phonostylistics, Sociophonology and Speech Acoustic Variability*, Florianópolis, Brazil.

Moraes, J. A. (1987). Correlates acoustiques de l'accent de mot Portugais Brésilien. *Proceeding of the the XI International Congress of Phonetic Sciences*, Tallin, Estônia.

Morlec, Y., Bailly, G. & Auberge, V. (1997). Generating the prosody of attitudes. *Proc. ESCA Workshop on Intonation: Theory, Models and Applications*. Athens, Greece, Sept 18-20, 1997, pp. 251-254.

Müller, K. (1992). Theatrical moments: on contextualizing funny and dramatic moods in the course of telling a story in conversation. In P. Auer & A. d. Luzio. *The Contextualization of Language*. Amsterdam, Benjamins: 199-221.

Mysak, E. D. & Hanley, T. D. (1958). Aging processes in speech: pitch and duration characteristics. *Journal of Gerontology* 13: 309-313.

Nakajima, S. & Allen, J. F. (1992). Prosody as a cue for discourse structure. *Proceeding of the the International Conference on Spoken Language Processing*, Banff, Canada.

Nakatani, C. (1993). Accenting on pronouns and proper names in spontaneous narratives. *Proceeding of the ESCA Workshop on Prosody*, Lund, Sweden.

Nakatani, C. H. & Hirschberg, J. (1995). Discourse structure in spoken language: Studies on speech corpora. *Proceeding of the AAAI Symposium Series: Empirical Methods in Discourse Interpretation and Generation.*

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht, Foris.

Nichols, P. C. (1989). Storytelling in Carolina: continuities and contrasts. *Anthropology and Education Quarterly* 20: 232-245.

Nittrouer, S., McGowan, R. S., Milenkovic, P. H. & Beehler, D. (1990). Acoustic measurements of men's and women's voices: a study of context effects and covariation. *Journal of Speech and Hearing Research* 33: 761-775.

Nooteboom, S. G. & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: an experiment. *Journal of the Acoustical Society of America* 82(5): 1512-1524.

O'Connell, D. C. & Kowal, S. (1972). Problems of measurement in psycholinguistic pause and rate research. *Proceeding of the XXth International Congress of Psychology*, Tokyo.

O'Connell, D. C. & Kowal, S. (1983). Pausology. In W. A. Sedelow & S. Y. Sedelow. *Computers in Language Research 2*. Berlin, New York, Amsterdam, Mouton Publishers. 19: 221-301.

O'Connell, D. C., Kowal, S. & Hörmann, H. (1969). Semantic determinants of pauses. *Psychologische Research* 2: 51-81.

O'Connell, D. C. & Kowal, T. D. (1972). Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research* 1: 155-164.

Ohala, J. & Ewan, W. (1973). Speed of pitch range. *Journal of the Acoustical society of America* 53: 345.

Oliveira, M. (1995). *A Narrativa Oral de Experiência Vicária*. Departmento de Letras. Recife, PE, Universidade Fedral de Pernambuco.

Oliveira, M. (1998). A comparative study between American English and Brazilian Portuguese intrinsic pitch of vowels. *Proceeding of the the 14th Northwest Linguistics Conference*, Vancouver, Simon Fraser University.

Oliveira, M. (1999a). The function of self-aggrandizement in storytelling. *Narrative Inquiry* 9 (1): 25-47.

Oliveira, M. (1999b). Narrative of Vicarious Experience Revisited. *Paper presented at the International Conference on Storytelling*, Brock University, St. Catherines.

Oliveira, M. (1999c). The role of pause in spontaneous narratives. *Proceeding of the 15th Northwest Linguistics Conference*, Victoria, University of Victoria.

Olson, W. C. & Koetzle, V. S. (1936). Amount and rate of talking of young children. *Journal of Experimental Education* 5: 175-179.

Ong, W. J. (1982). Oral Remembering and Narrative Structures. In D. Tannen. *Analyzing Discourse: Text and Talk*. Washington, D.C., Georgetown University Press: 12- 24.

O'Shaughnessy, D. (1976). *Modeling fundamental frequency, and its relationship to syntax, semantics and phonetics*, MIT.

O'Shaughnessy, D. (1992). Recognition of hesitations in spontaneous speech. *Proceeding of the IEEE.*

Osser, H. & Peng, F. (1964). A cross-cultural study of speech rate. *Language and Speech* 7: 120-125.

Ostendorf, M. (1993). Prosody. *Proceeding of the ARPA Human Language Technology Workshop*, San Francisco, Morgan Kaufmann.

Ostendorf, M., Price, P., Bear, J. & Wightman, C. W. (1990). The use of relative duration in syntactic disambiguation. *Proceeding of the Speech and Natural Language Workshop*, Hidden Valley, PA, DARPA.

Pakosz, M. & Flashner, V. (1988). Prosodic features and narrative strategies. *Papers and studies in contrastive linguistics* 24: 33-46.

Palmer, H. E. (1922). *English Intonation, with Systematic Exercises*. Cambridge, Heffer.

Parmenter, C. E. & Treviño, S. N. (1935). The length of the sounds of a Middle Westener. *American Speech* 10: 129-133.

Passonneau, R. J. & Litman, D. J. (1993). Intention-based segmentation: Human reliability and correlation with linguistic cues. *Proceeding of the 31st Annual Meeting of the Association for Computational Linguistics (ACL-93)*, Columbus, Ohio.

Passonneau, R. J. & Litman, D. J. (1997). Discourse Segmentation by Human and Automated Means. *Computational Linguistics* 23(1): 103-139.

Patterson, D. & Ladd, D. R. (1999). Pitch range modeling: linguistic dimensions of variation. *Proceeding of the ICPhS*, San Francisco.

Pawley, A. & Syder, F. (1977). The one clause at a time hypothesis. *Ms.*

Petersen, N. R. (1978). The influence of aspiration on the fundamental frequency of the following vowel in Danish: some preliminary observations. *ARIPUC* 12: 91-112.

Petersen, N. R. (1986). Perceptual compensation for segmentally conditioned fundamental frequency perturbation. *Phonetica* 43: 31-42.

Peterson, C. & McCabe, A. (1983). *Developmental Psycholinguistics: Three Ways of looking at a Child's Narrative*. New York, Plenum Press.

Peterson, G. E. & Barney, H. L. (1952). Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24(2): 175-184.

Phillips, J. (1970). *Formal characteristics of speech which mothers address to their young children*, John Hopkins University.

Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *Journal of the Acoustical Society of America* 66: 363-369.

Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*, MIT.

Pierrehumbert, J. & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan & M. Pollack. *Intentions in Communication*. London, The MIT Press: 271-311.

Pijper, J. R. d. (1983). *Modeling British English Intonation*. Dordrecht, Foris.

Pijper, J. R. d. & Sanderman, A. A. (1994). On the perceptual strength of prosodic boundaries and its relation to supresagmental cues. *J. Acoust. Soc. Am.* 96(4): 2037-2047.

Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor, University of Michigan.

Polanyi, L. (1979). So what's the point? *Semiotica* 25: 207-241.

Polanyi, L. (1981a). On telling the same story twice. *Text* 1: 315-336.

Polanyi, L. (1981b). What stories can tell us about their tellers' world. *Poetics Today* 2: 97-112.

Polanyi, L. (1982). Linguistic and social constraints on storytelling. *Journal of Pragmatics* 6: 509-524.

Polanyi, L. (1985a). Conversational Storytelling. In T. v. Djik. *Handbook of Discourse Analysis*. London, Academic Press. 3: 183-201.

Polanyi, L. (1985b). *Telling the American Story: From the Structure of Linguistic Texts to the Grammar of a Culture*. Norwood, Ablex.

Polanyi, L. (1988). A formal model of discourse structure. *Journal of Pragmatics* 12: 601-638.

Polkinghorne, D. E. (1988). *Narrative Knowing and the Human Sciences*. Albany, NY, State University of New York Press.

Pratt, M. L. (1977). *Towards a Speech Act Theory of Literary Analysis*. Bloomington, IN, Indiana University Press.

Prince, E. (1981). Toward a taxonomy of given-new distinction. In P. Cole. *Radical Pragmatics*. Cambridge, Academic Press: 223-256.

Prince, G. (1982). *Narratology: The Form and Functioning of Narrative*. Berlin, New York, Amsterdam, Mouton.

Propp, V. (1958). *Morphology of the Russian Folktale*. Austin, Texas, University of Texas Press.

Quirk, R. & al., e. (1985). *A Comprehensive Grammar of the English Language*. London, New York, Longman.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1972). *A Grammar of Contemporary English*. London, Longman.

Quirk, R., Svartvik, J., Duckworth, A. P., Rusiecki, J. P. L. & Colin, A. J. T. (1968). Studies in the correspondence of prosodic to grammatical features in English. In R. Quirk. *Essays in English Language: Medieval and Modern*. London, Longman: 120-135.

Reichman, R. (1985). *Getting Computers to Talk Like You and Me: Discourse Context, Focus, and Semantics*. Cambridge, Bradford.

Reinhart, T. (1981). *Pragmatics and Linguistics: An Analysis of Sentence Topics*. Bloomington, IULC.

Reinhart, T. (1984). Principles of Gestalt perception in the temporal organization of narrative texts. *Linguistics* 22: 779-809.

Reynolds, A. & Paivio, A. (1968). Cognitive and emotional determinants of speech. *Canadian Journal of Psychology* 22: 164-175.

Rietveld, A.C.M. & Gussenhoven, C. (1987). Perceived speech rate and intonation. *Journal of Phonetics* 15: 273-285.

Robins, R. H. (1972). Aspects of prosodic analysis. In V. B. Makkai. *Phonological Theory. Evolution and Current Practice*. New York, Holt: 264-274.

Robinson, J. A. (1981). Personal narratives reconsidered. *Journal of American Folklore* 94: 58-85.

Rochester, S. R., Thurston, S. & Rupp, J. (1977). Hesitation as clues to failures in coherence: A study of the thought-disordered speaker. In S. Rosenberg. *Sentence*

Production: Developments in Theory and Research. Hillsdale, N.J., Lawrence Erlbaum Associates.

Rooth, M. (1985). *Association with Focus.* Amherst, MA, University of Massachusetts.

Rosen, H. (1988). The autobiographical impulse. In D. Tannen. *Linguistics in context: connecting observation and understanding.* Norwood, NJ, Ablex: 69-88.

Rosenfield, B. (1987). *Pauses in Oral and written Narratives.* Boston, Boston University.

Rossi, M., Cristo, A. D., Hirst, D. J., Martin, P. & Nishinuma, Y. (1981). *L'Intonation: De L'Acoustique à La Sémantique.* Paris, Klincksieck.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. B. A. Collins. *Representation and Understanding: Studies in Cognitive Science.* New York, Academic Press.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In D. L. Berge & J. Samuels. *Basic processes in reading comprehension.* Hillsdale, NJ, Lawrence Erlbaum: 265-303.

Rumelhart, D. E. (1980). On evaluating story grammars. *Cognitive Science* 4: 313-316.

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

Sá, P., Cunha, D., Lima, A. & Oliveira, M., Eds. (1997). *A Linguagem Falada Culta na Cidade do Recife: Diálogo Entre Informante e Documentador.* Recife, Editora Universitária.

Sabin, E. J. (1976). *Pause and Rate Phenomena in Adult Narratives.* Science (Research). Saint Louis, Saint Louis University: 78.

Sabin, E. J., Clemmer, E. J., O'Connell, D. C. & Kowal, S. (1979). A pausological approach to speech development. In A. W. Siegman & S. Feldstein. *Of Time and Speech.* Hillsdale, New Jersey, Lawrence Erlbaum.

Sacks, H. (1971). Lecture Notes: Stories in Conversation. In G. Jefferson. *Lectures in Conversation.* Oxford, UK ; Cambridge, Mass., Blackwell.

Sacks, H. (1972). On the analyzability of stories by children. In J. J. Gumperz & D. Hymes. *Directions in Sociolinguistics.* New York, Holt, Rinehart, and Winston: 325-345.

Sacks, H., Schegloff, E. A. & Jefferson, G. (1977). The preference for self-cirrection in the organisation of repair in conversation. *Language* 53: 361-382.

Sagisaka, Y., Campbell, N. & Higuchi, N., Eds. (1997). *Computing Prosody: Computational Models for Processing Spontaneous Speech.* New York, Springer.

Sajavaara, K. & Lehtonen, J. (1980). The analysis of cross-language communication: prolegomena to the theory and methodology. In H. Dechert & M. Raupach. *Towards a Cross-Linguistic Assessment of Speech Production.* Frankfurt, Lang.

Sanderman, A. A. & Collier, R. (1997). Prosodic phrasing comprehension. *Language and Speech* 40(4): 391-409.

Schaffer, D. (1984). The role of intonation as a cue to topic management in conversation. *Journal of Phonetics* 12: 327-344.

Schank, R. & Ableson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ, Lawrence Erlbaum Associates.

Schegloff, E. A. (1979). The relevance of repair for syntax-for-conversation. In T. Givon. *Discourse and Syntax*. New York, Academic Press. 12: 261-286.

Schiffrin, D. (1981). Tense Variation in Narrative. *Language* 57(1): 45-62.

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

Schiffrin, D. (1994). How a story says what it means and does. *Text* 4: 313-346.

Schmerling, S. F. (1976). *Aspects of English Sentence Stress*. Cambridge, MA, MIT Press.

Schuetze-Coburn, S., Shapley, M. & Weber, E. G. (1991). Units of intonation in discourse: a comparison of acoustic and auditory analyses. *Language and Speech* 34(3): 207-234.

Schwitalla, J. (1992). Comments on: Margret Selting: Intonation as a contextualization device. In P. Auer & A. d. Luzio. *The Contextualization of Language*. Amsterdam, Benjamins.

Scollon, R. (1981). *Tempo, Density, and Silence: Rhythms in Ordinary Talk*. Center for Cross-Cultural Studies. Fairbanks, University of Alaska.

Scollon, R. (1982). The rhythmic integration of ordinary talk. In D. Tannen. *Analysing Discourse: Text and Talk*. Washington DC, Georgetown University Press: 335-349.

Scollon, R. & Scollon, S. B. K. (1981). *Narrative, Literacy and Face in Interethnic Communication*. Norwood, N. J., Ablex.

Selkirk, E. O. (1980). On prosodic structure and its relation to syntactic structure. *Indiana University Linguistics Club*.

Selkirk, E. O. (1984). *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge, Mass., MIT Press.

Selting, M. (1992). Intonation as a contextualization device: case studies on the role of prosody, especially intonation in contextualizing story telling in conversation. In P. Auer & A. d. Luzio. *The Contextualization of Language*. Amsterdam, Benjamins.

Selting, M. (1992). Prosody in conversational questions. *Journal of Pragmatics* 17: 315-345.

Selting, M. (1996). On the interplay of syntax and prosody in the constitution of turn-constructional units and turns in conversation. *Pragmatics* 6(3): 371-388.

Shapley, M. (1989). *Fundamental Frequency Variation in Conversational Discourse*. Department of Linguistics. Los Angeles, University of California: 210.

Shipp, T. & Hollien, H. (1972). Speech frequency and duration measures as a function of chronologic age. *Journal of the Acoustical Society of America* 51: 111.

Shriberg, E. E. & Lickley, R. J. (1993). Intonation of clause-internal filled pauses. *Phonetica* 50: 172-179.

Shriberg, E., Stolcke, A., Hakkani-Tur, D. & Tur, G. (2000), Prosody-Based Automatic Segmentation of Speech into Sentences and Topics. *Speech Communication* (Special Issue on Accessing Information in Spoken Audio) 32: 127-154.

Shuman, A. (1986). Storytelling Rights. The Uses of Oral and Written Texts by Urban Adolescents. Cambridge, Cambridge University Press.

Siegman, A. & Feldestein, S., Eds. (1979). Of Time and Speech. Hillsdale, N. J., Lawrence Erlbaum.

Siegman, A. W. (1979). Cognition and hesitation in speech. In A. Siegman & S. Felstein. *Of Time and Speech*. Hillsdale, NJ, Lawrence Erlbaum: 151-178.

Siegman, A. W. & Pope, B. (1966). Ambiguity and verbal fluency in the TAT. *Journal of Consulting Psychology* 30: 239-245.

Silverman, K. (1987). *Natural prosody for synthetic speech*. Cambridge, Cambridge University.

Sluijter, A. M. C. & Terken, J. M. B. (1993). Beyond sentence prosody: paragraph intonation in Dutch. *Phonetica* 50: 180-188.

Smith, V. L. & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language* 32: 25-38.

Snell, A. L. F. (1918). Contributions to rhetorical theory. In F. N. Scott. *Pause: A Study of its Nature and its Rhythmical Function in Verse, Especially Blank Verse*. Ann Arbor, Michigan, The Ann Arbor Press. 8.

Snidecor, J. C. (1943). A comparative study of the pitch and duration characteristics of impromptu speaking and oral reading. *Speech Monographs* 10: 50-56.

Song, F. & Cohen, R. (1991). Tense interpretation in the context of narrative. *Proceeding of the AAAI*.

Starkweather, C. W. (1983). *Speech and language : principles and processes of behavior change*. Englewood Cliffs, N.J., Prentice-Hall.

Steedman, M. (1990). Structure and intonation in spoken language understanding. *Proceeding of the 28th Annual Meeting of the Association for Computational Linguistics*, Pittsburgh.

Steele, S. A. (1986). Interaction of vowel Fo and Prosody. *Phonetica* 43: 92-105.

Steer, A. B. (1974). Sex differences, extraversion and neuroticism in relation to speech rate during the expression of emotion. *Language and Speech* 17: 80-86.

Stein, N. L. & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle. *New Directions in Discourse Processing*. Norwood, NJ, Ablex. 2.

Stenström, A.-B. (1986). A study of pauses as demarcators in discourse and syntax. In J. Aarts. *Corpus Linguistics II: new studies in the analysis and exploration of computer corpora*. Amsterdam, Rodopi: 203-217.

Stenström, A.-B. (1990). Pauses in monologue and dialogue. In E. Svartvik. *The London-Lund Corpus of Spoken English. Description and Research*. Lund, Lund University Press: 211-252.

Stifelman, L. J. (1993). User Repairs of Speech Recognition Errors:

An Intonational Analysis. *MIT Media Laboratory Technical Report*.

Stifelman, L. J. (1995). A discourse analysis approach to structured speech. *Proceeding of the AAAI*, Stanford.

Strangert, E. (1991). Where do pauses occur in texts read aloud? *Proceeding of the 12th Scandinavian Conference of Linguistics*, Reykjavik.

Strangert, E. (1993). Speaking style and pausing. *PHONUM* 2: 121-137.

Streeter, L. A. (1978). Acoustic determinants of phrase boundary location. *Journal of the Acoustical Society of America* 64: 1582-1592.

Svartvik, J. & Quirk, R. (1980). *A Corpus of English Conversation*. Lund, Sweden, CWK Gleerup.

Svetozarova, N. D. (1975). The inner structure of intonation contours in Russian. In G. Fant & M. A. A. Tatham. *Automatic Analysis and Perception of Speech*. London, Academic.

Sweet, H. (1906). *A Primer of Phonetics*. Oxford, Clarendon Press.

Swerts, M. (1994). *Prosodic Features of Discourse Units*. Eindhoven, Eindhoven University of Technology.

Swerts, M. (1997). Prosodic features at discourse boundaries of different strenght. *Journal of the Acoustical Society of America* 101(1): 514-521.

Swerts, M., Bouwhuis, D. G. & Collier, R. (1994). Melodic cues to the perceived "finality" of utterances. *J. Acoust. Soc. Am.* 96(4): 2064-2075.

Swerts, M. & Collier, R. (1992). On the controlled elicitation of spontaneous speech. *Speech Communication* 11(4-5): 463-468.

Swerts, M., Collier, R. & Terken, J. (1994). Prosodic predictors of discourse finality in spontaneous monologues. *Speech Communication* 15: 79-90.

Swerts, M. & Geluykens, R. (1993). The prosody of information units in spontaneous monologue. *Phonetica* 50: 189-196.

Swerts, M. & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech* 37: 21-43.

Swerts, M., Geluykens, R. & Terken, J. (1992). Prosodic correlates of discourse units in spontaneous speech. *Proceeding of the International Conference on Spoken Language Processing*, Banff.

Swerts, M., Strangert, E. & Heldner, M. (1996). Fo declination in read-aloud and spontaneous speech. *Proceeding of the International Conference on Spoken Language Processing*, Philadelphia.

Swerts, M., Wichmann, A. & Beun, R.-J. (1996). Filled pauses as markers of discourse structure. *Proceeding of the International Conference on Spoken Language Processing*, Philadelphia.

Szawara, J. & O'Connell, D. C. (1977). Temporal reflections of spontaneity in homilies. *Bulletin of the Psychonomic Society* 9: 360-362.

't Hart, J. & Collier, R. (1975). Integrating different levels of intonation analysis. *Journal of Phonetics* 3: 235-255.

't Hart, J., Collier, R. & Cohen, A. (1990). *A Perceptual Study of Intonation*. Cambridge, Cambridge University Press.

Tannen, D. (1979). What's in a frame? In R. Freedle. *New Directions in Discourse processing*. Norwood, Ablex: 137-181.

Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language* 58: 1-21.

Tannen, D. (1983). I take out the Rock-Dok! How Greek women tell about being molested (and create involvement). *Anthropological Linguistics* 25: 359-374.

Tannen, D. (1985). Relative focus on involvement in oral and written discourse. In D. R. Olson, N. Torrance & A. Hildyard. *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing*. Cambridge, Cambridge University Press: 124-147.

Tannen, D. (1989). *Talking Voices. Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge, Cambridge University Press.

Tannenbaum, P. H., Williams, F. & Hillier, C. S. (1965). Word predictability in the environment of hesitations. *Journal of Verbal Learning and Verbal Behaviour* 4(134-140).

Taylor, G. (1986). The development of style in children's fictional narrative. In A. Wilkinson. *The Writing of Writing*. Milton Keynes, Open University Press: 215-33.

Taylor, I. (1969). Content and structure in sentence production. *Journal of Verbal Learning and Verbal Behavior* 8: 170-175.

Tedlock, D. (1972). *Finding the Center: Narrative Poetry of the Zuni Indians*. New York, Dial Press.

Terken, J. (1991). Fundamental frequency and perceived prominence of accented syllables. *J. Acoust. Soc. Am.* 89(4).

Terken, J. M. B. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech* 27(3): 269-289.

Thorndyke, P. W. (1977). Cognitive structure in comprehension and memory of narrative discourse. *Cognitive Psychology* 9: 77-110.

Thorsen, N. (1979). Interpreting raw fundamental-frequency tracings of Danish. *Phonetica* 36: 57-78.

Thorsen, N. (1980). Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish. *Annual Report of the Institute of Phonetics, University of Copenhagen* 14: 121-205.

Thorsen, N. G. (1985). Intonation and text in Standard Danish. *Journal of the Acoustical Society of America* 77: 1205-1216.

Tielen, M. T. J. (1992). *Male and Female Speech. An Experimental Study of Sex-related Voice and Pronunciation Characteristics*, University of Amsterdam.

Todorov, T. (1969). Structural analysis of narrative. *Novel* 3: 23-49.

Todorov, T. (1970). Les Transformations Narratives. *Poétique* 3: 322-333.

Todorov, T. (1977). *The Poetics of Prose*. Oxford, Blackwell.

Toolan, M. J. (1988). *Narrative: A Critical Linguistic Introduction*. London, Routledge.

Tosi, O. (1965). *A method for acoustic segmentation of continuous sound into pauses and signals and measurement of segment durations*. Ohio, Ohio State University.

Uhmann, S. (1992). Contextualizing relevance: on some forms and functions of speech rate changes in everyday conversation. In P. Auer & A. d. Luzio. *The Contextualization of Language*. Amsterdam, Benjamins. 297-336.

Uldall, E. (1960). Attitudinal meanings conveyed by intonation contours. *Language and Speech* 3: 223-234.

Uldall, E. (1964). Dimensions of meaning in intonation. In D. A. e. al. *in Honor of Daniel Jones: Papers contributed on the occasion of his 80th birthday*. London, Longman: 271-279.

Umeda, N. (1981). Influence of segmental factors on fundamental frequency in fluent speech. *J. Acoust. Soc. Am.* 70(2): 350-355.

Umeda, N. (1982). Fo declination is situation dependent. *Journal of Phonetics* 10: 279-290.

Umiker-Sebeok, D. S. (1979). Preschool children's intraconversational narratives. *Journal of Child Language* 6: 91-109.

Van De Water, D. A., Monti, L. A., Kirchner, P. B. & O'Connell, D. C. (1987). Speaking and writing: Comparisons of two psycholinguistic siblings. *Bulletin of the Psychonomic Society* 25: 99-102.

Van Dijk, T. A. (1972). *Some Aspects of Text Grammars. A Study in Theoretical Linguistics and Poetics*. The Hague, Mouton.

Van Dijk, T. A. (1984a). *News as Discourse*. Hillsdale, NJ, Lawrence Erlbaum.

Van Dijk, T. A. (1984b). *Prejudice in Discourse: An Analysis of Ethnic Prejudice in Cognition and Conversation*. Amsterdam, John Benjamins.

Van Dijk, T. A. (1988). *News Analysis*. Hillsdale, NJ, Lawrence Erlbaum.

Van Donzel, M. (1999). *Prosodic Aspects of Information Structure in Discourse.* Faculteit der Geesteswetenschappen. Amsterdam, University van Amsterdam: 195.

Van Donzel, M. E. (1994). How to specify focus without using acoustic features. *Proceedings of the Institute of Phonetic Sciences*(18): 1-17.

Van Heuven, V. J. (1994). What is the smallest prosodic domain? In P. A. Keating. *Phonological structure and phonetic form, Papers in laboratory phonology III.* Cambridge, Cambridge University Press: 76-98.

Venditti, J. J. & Swerts, M. (1996). Intonational cues to discourse structure in Japanese. *Proceeding of the ICSLP.*

Verzeano, M. & Finesinger, J. E. (1949). An automatic analyzer for the study of speech in interaction and in free association. *Science* 110: 45.

Vroomen, J., Collier, R. & Mozziconacci, S. (1993). Duration and intonation in emotional speech. *Proceeding of the Eurospeech*, Berlin, Germany.

Walker, A. G. (1982). Discourse rights of witnesses: Their circumscription in trial. *Sociolinguistic Working Paper* 95.

Wallin, J. E. W. (1901). Researches in the rhythm of speech. *Studies from the Yale Psychological Laboratory* 12: 54-67.

Wang, M. Q. & Hirschberg, J. (1992). Automatic classification of intonational phrase boundaries. *Computer Speech and Language* 6: 175-196.

Watson, K. A. (1972). *The rhetoric of narrative structure: A sociolinguistic analysis of stories told by part-Hawaiian children*, University of Hawaii.

Webber, B. (1988). Discourse deixes: Reference to discourse segments. *Proceeding of the The 26th Annual Meeting of the Association for Computational Linguistics*, Buffalo.

Whalen, D. H. & Levitt, A. G. (1995). The universality of intrinsic Fo of vowels. *Journal of Phonetics* 23: 349-366.

Wichman, A., House, J. & Rietveld, T. (1997). Peak displacement and topic structure. *Proceeding of the ESCA Tutorial and Research Workshop on Intonation*, Athens.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *J. Acoust. Soc. Am.* 91(3): 1707-1717.

Wilkes, A. L. & Kennedy, R. A. (1969). Relationship between pausing and retrieval latency in sentences of varying grammatical form. *Journal of Experimental Psychology* 79: 241-245.

Wilkinson, J. (1986). Describing Children's writing: Text evaluation and teaching strategies. In J. Harris & J. Wilkinson. *Reading Children's Writing: A Linguistic View.* London, George Allen and Unwin: 11-31.

Willems, N. (1982). *English Intonation from a Dutch Point of View.* Dordrecht, Foris.

Windsor, J. & Fristoe, M. (1991). Key word signing: Perceived and acoustic differences between signed and spoken narratives. *Journal of Speech and Hearing Research* 34: 260-268.

Wolfson, N. (1976). Speech events and natural speech. *Language in Society* 5: 189-209.

Wolfson, N. (1979). The conversational historical present alternation. *Language* 55(1): 168-182.

Wolfson, N. (1982). *CHP: The Conversational Historic Present in American English Narrative*. Dordrecht, Foris.

Wood, S. (1975). Speech tempo. *Working Papers of the Phonetics Laboratory, Lund University* 9: 99-147.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24: 245-261.

Yule, G. (1980). Speakers' topics and major paratones. *Lingua* 52: 33-47.

Zhang, X. & Sang, Y. (1986). *Chinese Profiles*. Beijing, Panda.