

# Analyzing Narrative Informativeness in Speech and Writing\*

*Ruth A. Berman and Dorit Ravid*

## 1. Introduction

The topic of organization of information has been considered in functional linguistics at the level of both sentence and extended discourse. At sentence level, various approaches have been proposed to account for such contrasts as old versus new, given versus non-given, topical versus nontopical, activated versus nonactivated information (e.g., Lambrecht 1994; Prince 1981; Vallduví and Engdahl 1996). Our study is more closely aligned with research on extended discourse (e.g., Biber, Conrad, and Reppen 1998; Chafe 1994; Halliday 1989). Like these scholars, we view the “discourse event” as being shaped by the interaction between the two dimensions of modality (speech and writing) and genre (narrative and non-narrative). To these, we add a *participant* dimension by comparing the information packaging in texts produced by gradeschool children, adolescents, and adults. Our study also differs from much other research by using as its data base the “raw”, unedited texts elicited from non-expert native speaker-writers as recorded and transcribed in speech and writing (Section 3.1).<sup>1</sup>

We compare spoken with written texts, on the assumption that the flow of information in discourse is modality-dependent and that different pressures of on-line processing apply in speech compared with displaced time in writing, and that these distinctions lead to information being processed and hence presented differently in the two modalities (Chafe 1994; Strömqvist, Nordqvist, and Wengelin 2004). In order to focus on these facets of modality-driven differences, the present analysis disregards other important facets of writing versus speech. In what follows, we thus avoid such issues as: writing as a notational system and concern with features of alphabetization (Tolchinsky 2003) or spelling (Pacton and Fayol 2004); the nature of written language “as a special discourse style” (Ravid and Tolchinsky 2002); and the more complex forms of linguistic expression and higher level of usage associated with written compared with spoken texts (Malvern et al., 2004; Berman, Nir-Sagiv, and Bar-Ilan in press; Ravid

2004). Rather, we consider here the effects of processing constraints on the flow of information in narrative text construction, with the aim of demonstrating that texts produced in speech are less carefully monitored and show more effects of the pressures of online production of linguistic output in speaking compared with the offline activity of writing.

Underlying the analysis are several assumptions. First, we propose that the way in which information is presented in extended discourse can be treated as distinct from (although not unrelated to) linguistic means used for achieving discourse connectivity (Berman 1998), “clause combining” (Haiman and Thompson 1988) or “syntactic packaging” (Berman and Slobin 1994: 538–554). Second, this idea of informativeness is also distinct from, although not independent of, thematic or referential content. That is, in principle, the same type of analysis should apply to the sub-genre of narratives investigated here (personal-experience narratives dealing with interpersonal conflict) and to other narrative subgenres such as adventures or mysteries. But since we view discourse information as essentially *genre-dependent*, rather different principles might apply to the categories of information characterizing other types of discourse such as conversation, description, or expository texts.

Third, in developmental perspective, we assume that even nine-year-olds (the youngest group in our study) make some distinction between texts they produce in the two modalities – although it has been shown that their online processing of written texts is still more locally confined than that of older children (Wengelin and Strömquist 2004). We also assume that the texts written by gradeschool children are less differentiated from their spoken counterparts in level of usage and linguistic register than those of older children – especially from high school up (Bar-Ilan and Berman in press; Jisa 2004).

The goal of this study is to present an empirically anchored, text-based model of “information parsing” that will account for differences in text construction along the variables of modality, genre, and development. To this end, we propose criteria for characterizing narrative information packaging along the two dimensions of level of informativeness and types of informative material in narrative discourse as illustrated by personal-experience narratives produced in speech and writing by children, adolescents, and adults (Section 2). This is followed by description of findings from initial application of these criteria to narratives produced by speaker-writers of English and Hebrew (Section 3), concluding with a discussion of the cross-modal, cross-linguistic, and developmental implications of these preliminary findings (Section 4).

## 2. Informativeness in narratives: A proposed analysis

The notion “level of informativeness” derives from the distinction between two major components of textual material: informative – defined as *novel contentive material* (Section 2.1) and *non-informative material*, which includes both novel but redundant and non-novel or extraneous material (Section 2.2).

### 2.1 Eventive, Descriptive, and Interpretive Information Units

The basic element of analysis for what we count as “informative material” is an *information unit* or “Infu”, which by definition must contain novel information. Since Infus provide novel contentive material, they cannot simply reiterate or reformulate previous information. Semantically, Infus correspond to discourse-functional units comparable to Chafe’s (1994) “idea units”; syntactically, they may be less than a clause, but are typically one or more clauses long.

In discussing facets of narrative content, different, though largely complementary, perspectives have been adopted on issues relating to “referential” versus “evaluative” elements as proposed by Labov (1972; 1997) – as re-evaluated and re-analyzed by Aisenman and Assayag (1999), Reinhart (1995), Segal (2001) – or background states versus foreground events – as considered in Berman and Slobin (1994), Hopper (1979), Reinhart (1984). Here, following Berman (1997), we distinguish three types of informative material in narratives – eventive, descriptive, and interpretive. *Eventive Infus* are similar to Labov’s (1972) “narrative” or referential clauses; they encode plot-advancing, typically sequentially ordered events, and serve to anchor descriptive and interpretive Infus within a narrative frame. *Descriptive Infus* provide factual information on the circumstances surrounding these events; they make reference to states of affairs and motivations that provide the background to the events that constitute the story. *Interpretive Infus* are closer to canonic notions of narrative evaluation; they express narrators’ subjective perspective on the events recounted and their interpretations of the attitudes or internal states which they attribute to participants in these events.

Below, we define and motivate these three types of novel informative material on the basis of narratives produced by English and Hebrew speakers of different ages who were asked to write and tell a story about an incident in which they had been involved in interpersonal conflict (see Sec-

tion 3.1). As our basic unit of text division we use the *clause*, in the sense of “a unified predication”, following Berman and Slobin (1994: 660-664). In contrast, an Infu is a unit of discourse rather than a semantic proposition or a syntactic construction, and so it can be defined both within and beyond the boundaries of a single clause.

Different types of Infus occur in the text in (1), the written narrative produced by an 11<sup>th</sup> grade boy at a Californian high school who had been asked to tell and write a story about an incident where he had been involved in interpersonal conflict. Clauses are numbered consecutively, embedded clauses are marked in angle brackets, and material defined as Eventive Infus is marked in bold.<sup>2</sup>

(1) **Written Narrative of High School Boy - Grade XI**

1) *When I was in the seventh grade*, 2) *I had a conflict with a boy* 3) *who was in a few of my classes*. 4) *As it turned out*, 5) *his father was an executive vice-president at the company* 6) *where my father worked*. 7) *The boy was constantly giving me grief*, 8) *saying that* 9) *if I ever did anything* 10) *to upset him*, 11) *he would have* 12) *my father fired*. 13) *I knew* 14) *this was ridiculous*, 15) *but nevertheless it was plenty annoying*. 16) *The boy was not just annoying to me*, 17) *he had conflicts with at least ten other people* 18) *I knew*, 19) *not exaggerating*.

20) ***So one day we went to the counseling office at the school***. 21) ***The counselor told us*** 22) *that* <23) *since the teachers had not reported anything*>, 22) *we had no proof of the boy's actions*. 24) ***So the administration at the school did nothing***. 25) ***I visited the principal***, 26) ***but he did not take any action either***, 27) *so the boy kept up his incessant pestering*. 28) ***And one day I finally snapped***. 29) *When the teacher was out of the room*, 30) ***the boy said something to me***, 31) ***and I turned around*** 32) ***and confronted him***. 33) *The boy told me* 34) *that I would not* 35) *or could not do anything to him*, 36) *so I proved him wrong*. 37) ***I hit him***. 38) *And from that day on he stayed away from me*. 39) *I probably should not have resorted to that action*, 40) *but nothing else* 41) *I had done* 40) *worked*.

The text in (1) starts out largely with what we define as Descriptives – as motivated below. The first *Eventive* Infus occur in clauses #20 - #21, starting canonically with “So one day” (Berman 2001), followed by other *Eventives* in Clauses #25 “I visited the principal” and #28 “I finally

snapped”. These illustrate canonical narrative clauses, chronologically sequential and temporally specific. As such, these Eventives constitute the episode that forms the background to the confrontation between the narrator and his antagonist in Clauses #31 to #33 and the narrator’s action in Clause #37 “I hit him”.

Bear in mind that Information Units – that is, narrative material that constitutes novel information – cannot be unequivocally identified by semantic content or syntactic form alone. Infus of the type we term “Eventives” are relatively easy to identify; but drawing the line between Eventives and Descriptives, and specifying the material that we analyzed as Descriptive (and hence non-Eventive) proved to be rather less straightforward. Issues that we addressed in demarcating different types of Information Units involve both discourse functions and linguistic features, including: background events; reported and direct speech; negation; use of habitual, iterative, and protracted aspect; and stative predicates. In what follows, we outline our decisions in defining such material as either Eventives or Descriptives.

Background Eventives – like clauses #20-21 in (1) above – represent material that is clearly dynamic and plot-advancing, hence “eventive”, even though it refers to background events, the setting (Berman 2001), or orientation (Labov 1972) to the events making up the episodes that constitute the story. We therefore treated such material as a subclass of Eventives.

Reported and direct speech: Another issue in defining what constitutes “a plot-advancing sequential event” concerns complement clauses, particularly in cases of reported and direct speech. These are discussed in the literature as either ‘verbal representations’ or as ‘demonstrating representations’ (Clark and Gerrig 1990; Coulmas 1986). In our analysis, complement clauses that are introduced by speech act verbs like *told*, *demand* were coded as Descriptive while their introducing clauses were counted as Eventives. This is represented by the sequence in (1) “The counselor told us [EVENTIVE] that we had no proof of the boy’s actions” [DESCRIPTIVE]. Thus, in (2), an excerpt from the oral story of a Hebrew-speaking man about a student of his named Paul in biology class, the bolded material represents Eventives while all the rest are Descriptive.

(2) **Excerpt from Story Told in Hebrew by Adult**<sup>3</sup>

*And Paul started to interrogate, and to ask where I found it, and where he could find some and how you get there and what not. ... [Later] Paul explained to them how he caught the mantises and his methods of hunting, and where he saw them, and how he saw*

*to food for them, and that they normally eat crickets and where he caught them and what he caught them with.*

Negative Eventives: In Labov’s original scheme, negative clauses were regarded as “evaluative” rather than referential narrative elements (and see, too, Bamberg and Damrad-Frye’s [1987] study of children’s narratives). However, following other studies that have queried this criterion (Aisenman and Assayag 1999; Berman and Slobin 1994, pp. 6-9; Segal 2001), we do not automatically treat negated propositions as non-Eventive. The fact that a protagonist refuses to do something, does not do something, or avoids doing something is quite often a plot-advancing element, or it may be the trigger for what happens next. For example, Clause #26 in the text in (1) “but he did not take any action either”, like the preceding negative in Clause #24 “The administration at the school did nothing” are both clearly events that trigger what the boys did in response, and so were analyzed as Eventives.

Aspectually Marked Descriptions: Predications in habitual, protracted, or iterative aspect cannot strictly speaking be counted as events since they are not temporally anchored in a specific time. We generally defined these as Descriptive, taking into account their discourse role in context rather than as an across-the-board grammatically determined criterion (Hopper 1979). In our analysis, Descriptive material includes grammatically inflected progressive and perfect forms (e.g., “the boy was constantly giving me grief”, “the teachers had not reported anything” in [1] above); aspectual verbs (e.g. “kept up his incessant pestering”); and also aspectual adverbials like *for a long time*, *all the time*, *over and over*, *continually* – particularly relevant to a language like Hebrew, that lacks grammaticized aspect. The bolded material in the excerpt in (3) illustrate such aspectually marked non-eventives.

- (3) Excerpt from Oral Story of Woman Graduate Student  
*For months he would write me letters, call me on the phone, and come over to my house. At parties he would follow me around and bug me. At school he would constantly try to talk to me. This was satisfaction. He had screwed up and he was going to pay.*

Stative Verbs, including *be* and *have* as main verbs, are treated as Descriptive background, e.g., in (1) “When I was in 7<sup>th</sup> grade, I had a conflict”, as are verbs referring to mental or affective states, e.g., “I know that I am not always cooperative” or “I really hated her for what she did”. Non-dynamic predicates like “have an argument” may, however, be treated as Eventive, as periphrastic versions of, say, “We argued about it”.

A combination of criteria were thus required to distinguish Eventive and Descriptive material in narratives. In contrast, *Interpretives* – the narrator’s subjective commentary and perspective on events – are quite clearly distinct from these other two types of Information Units. The high school text in (1), as noted, consists largely of Descriptives, with relatively few Eventives and even fewer Interpretive elements. These occur in the protagonist’s perception of his classmate’s threat to have his father fired as “(I knew) this was ridiculous, but ...”, and of his own behavior in response to the antagonist at the end of the story, when he says “I probably should not have resorted to that action”. These interpretive elements typically occur in, but are by no means confined to, the setting and coda of narratives as in (1). For example, towards the middle of her account, a woman narrator says of her co-worker antagonist that she “had a very volatile disposition”, and of herself that “I am not usually a very confrontational person and it takes a lot for me to get mad”.

We have gone into some detail about our decisions in demarcating Eventive compared with Descriptive material, while briefly illustrating what we interpreted as Interpretation. This appeared necessary, since the tripartite division into types of narrative information proposed by Berman (1997) on the basis of oral narratives produced by young children is applied here to both written as well as spoken narratives from adolescents and adults. Most critical for the present analysis is the fact that the three classes of narrative content we identify – Eventive, Descriptive, and Interpretive – represent different types of *Information Units*, which together make up the informative substance of narrative discourse and so contribute uniquely to its “informativeness”.

## 2.2 Non-Novel and Non-Informative material

The other major block of narrative content consists of what we define as “non-informative” material since it fails to encode *new* narrative information, eventive, descriptive, or interpretive. We identify three classes of non-novel material: (1) Contentive, but non-novel; (2) non-contentive; and (3)

extra-contentive. The first class of non-novel contentive material refers to reiterations of propositional content and to reformulations of information that has already been mentioned. These different types of non-informative material are bolded in the text in (4), the oral version of the narrative of the high school boy who had previously written the same story that is reproduced in (1) above.

(4) **Oral Version of Story Written by High School Boy – Grade XI**

1) **I guess** 2) *the first one that comes to mind*, 3) *since we were talking about my dad, I went to **school** middle school out in La Jolla.* 5) *And one of the kids in my class,* 6) *<it turned out>* 5) *his dad was my dad's boss,* 7) *not directly,* **but he** *his dad was a an executive vice president of the company,* 8) *and my dad was uh **one of the lab** one of the guys* 9) *that **works actually** works in the lab.* 10) **And so** *he was constantly uh* 11) *<I don't know>* 10) **just** *bugging me,* 12) *saying* 13) **“you know** *if you mess with me* 14) *my dad can have your dad* 15) *fired* **and all this stuff** “. 16) *I was like,* 17) **“Shut up!” you know.** 18) **And he, I don't know,** *he just eventually got on my nerves really bad.* 19) **And the teacher was out of the classroom** 20) *and he said something,* 21) *and I just turned around* 22) *and like uh **how can I say?*** 23) *I didn't really assault him,* 24) *but I like pushed him real hard* **you know,** 25) *and I was like* 26) **“Don't mess with me,** 27) *you know* 28) *it makes me mad* 29) *when you talk about my dad that way”.* 30) **And** *he's like* 31) **“Yeah, what are you going to do about it?”** 32) **So then I** *smacked him,* 33) *and I didn't get in trouble* 34) *because the teacher was out of the classroom.* 35) *But he never bothered me again.* 36) *So I got lucky.*

The bolded “non-informative” material in (4) accounts for over 20% of this text (51 out of 233 words). This is in marked contrast to its written counterpart in (1), where the only material that might be counted as lying outside the three types of Infus, that is, as non-referentially informative, is the qualifying expression “not exaggerating” in Clause #19.

The first class of “non-informatives” we identify are, in fact, contentive, but they are non-novel, since they *reiterate* information that has already been provided or will be provided later in the narrative – like “the teacher was out of the classroom” in Clause #19 and again in Clause #34 of (4).

The second class of such elements divides up into *disfluencies* and *discourse markers*. Disfluencies include false starts, repairs, and repetitions

– like the bolded items in Clauses #3 and #8 of (4).<sup>4</sup> These noncontentive elements, which form part of what Clark (1996) terms “collateral” material, typically arise due to the pressures of rapid, online processing of linguistic output in the course of producing spoken discourse.

The other “noncontentive” elements are items termed variously discourse markers (Fleischman and Yaguello 2003; Jucker and Ziv 1998; Schiffrin 1995), pragmatic markers (Brinton 1996), or *particules énonciatives* (Fernandez 1994). These take the form of non-reiterated lexical elements which, like repairs and other disfluencies, are not part of the referential contentive material of a given proposition. We identified four types of discourse marking elements: (i) “interactive” items that are sender- or addressee-oriented (Berman, Ragnarsdóttir, and Strömqvist 2002) and communicative in intent, e.g., *I guess, you know* in (4) above; (ii) “segment-taggers” at segment beginnings – e.g., *okay, well, yeah*, and also frequent use of utterance-initial *and, so* in the text in (4) – and at segment endings – *that’s about it*; (iii) “qualifiers” – both intensifiers like *a whole lot, really* – and hedges like *basically* or *just, like* in (4); and (iv) online “monitoring” remarks, e.g., *how can I put it, let’s say*.

In our analysis, these different types of discourse markers were counted by occurrence as tokens rather than as types. And care was taken to consider their use in the text, so that the term *like* would not be counted as a discourse marker in a context such as “and then she *went like / and I was like* ‘don’t mess with me’ ” since it is, rather, a slang variety of the speech act verb *say*.

The elements bolded in (4) are typical, although not confined to, oral discourse, and mark a distinct contrast between written texts produced by the same person on the same topic across the data-base. This difference is clear in comparing the high schooler’s written and spoken texts in (1) and (4). And these features of oral monologic text production are clearly noticeable in the bolded elements in the oral narrative of a 7<sup>th</sup> grade girl who first wrote and then told the same story.

(5) **Oral Story Told by Junior High School Girl – Grade VII**

*Um well, alright, okay we have this one friend, and and she's like really rude. And so because she didn't like our other friend, because she was jealous of her, because she was like better than her, so she didn't so she want so she didn't want us to be her friend. And so we um so she would kind of like exclude her, and so we just like "You can either be our friend and her friend or not be our friend". And so, well, she accepted her you know that, but so now she's like better friends with her. But there's still like some problems with them.*

A third and final class of non-informative material differs from the other categories discussed so far in this section (reiterations, disfluencies, and discourse markers) since it lies outside the story frame and so is analyzed as "story-external" material. This includes expository-like generalizations and various kinds of meta-textual and inter-textual commentary. This is illustrated by the end of the story written by the woman whose text is excerpted in (3) above, about her breakup with a longtime boyfriend.

(6) **Closing Part of Narrative Written by Woman**

*Trust is something that is earned over time, during which you do not get caught lying. This is something I live by today. Conflict can teach you a lot about people and about yourself. I learned a lot from that conflict.*

The closing sentence in (6) is a classic coda since it goes back to the story itself, hence is by definition not extraneous to it, and so would be classed as an Interpretive Infu. But the three sentences preceding it are story-external material in the form of generalized comments lying outside the story frame. These represent a sophisticated departure from the narrower framework of narrative action structure, and are typical of more maturely proficient storytellers.

To sum up this section, Figure 1 shows the conceptual structure of the model for analyzing our data, distinguishing two major components of Informative versus Non-Informative material.

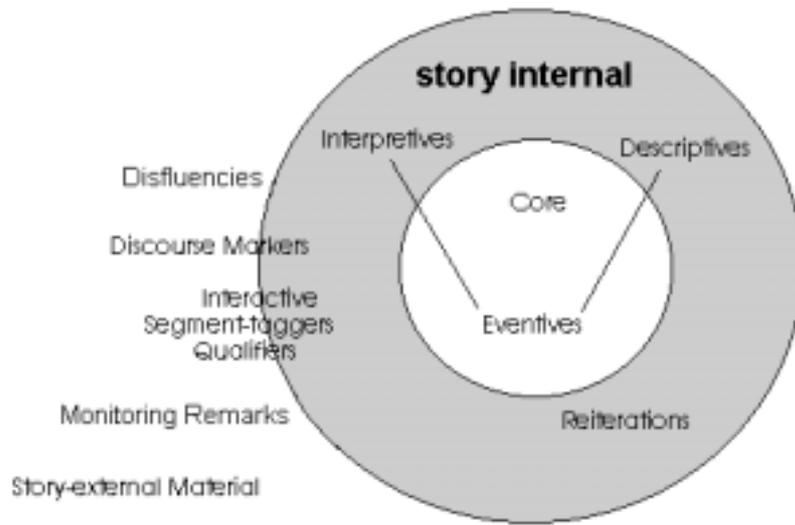


Figure 1. Informative and non-informative elements within, across, and beyond the story frame

Information Units – eventives, descriptives, and interpretives – frame the story, with eventives at the core of the narrative, anchoring descriptives and interpretives within a narrative frame. Reiterative non-informative units repeat the same narrative information, also within the narrative frame. Cutting across the narrative frame is non-informative, although communicative material (disfluencies and discourse markers). Monitoring remarks are more text-oriented and, together with story-external material, lie outside the narrative frame.

### 3. Preliminary analysis

The categories described in Section 2 were applied to spoken and written personal-experience narratives produced by schoolchildren, adolescents, and adults, native speakers of Californian English and Israeli Hebrew, who

were asked to tell and write a story on an incident of interpersonal conflict in which they had been involved.

### 3.1. Sample data-base

The data-base of the present study is taken from a subset of the Hebrew and the English-language sample of a large-scale cross-linguistic project in which closely comparable written and spoken texts were produced by schoolchildren and adults, native speakers of different languages (Berman 2005; Berman and Verhoeven 2002).<sup>5</sup> Subjects in seven countries, in the same four age-groups (grade schoolers aged 9–10 years, junior high schoolers aged 12–13, high school seniors aged 16–17, and graduate-level university students) were shown a three-minute wordless video clip depicting different conflict situations in a school setting. After seeing the video clip, each participant was required to produce four texts in randomly balanced order. They were asked to write and tell a story about an incident where they had been involved in a situation of “problems between people” (a personal-experience narrative) and to write a composition and give a talk in which they discuss the topic “problems between people” (an expository discussion). Thus both narrative and expository texts were elicited on the shared, socially relevant theme of interpersonal conflict. The vast majority of subjects across the different languages – around 90% from the youngest age-group up – followed these instructions, in the sense that they told a story in one case (typically in past tense and/or perfective aspect, with highly specific, personal reference to people, times, and places) and expressed general ideas (typically in the timeless present, with mainly impersonal or generic reference to people and situations) in the other.

The study reported here analyzes part of the written and spoken narrative texts produced in speech and writing by 26 different subjects – 4 in English and 4 in Hebrew at each of the three school-age groups (Grades IV, VII, XI) and one adult in each language – yielding a total of 52 narrative texts. These are all “authentic” texts, since they are unpublished, and were elicited from “naïve” speaker-writers without any editing or revision on our part. And they are elicited from speaker-writers of middle-class “standard language” background (Jisa 2004), where the adults were non-specialists in language and non-experts in writing but well-educated and “mainstream” in social class and literacy background.

### 3.2. Initial Trends for Narrative Information Packaging

The present study adopts a “case-study” approach, applying the model we have formulated to individual instances out of a far larger data-base, aimed at initial testing of its applicability.<sup>6</sup>

Our analysis involves the two variables of modality (speech versus writing) and age or level of literacy. We predicted, first, that written texts would reveal denser information packaging than their spoken counterparts, where informational density is defined as proportion of informational, “contentive” material compared with extraneous “non-informative” material. Second, this proportion would change as a function of age and increased literacy, and there would be less difference in level of informativeness between the written and spoken texts of the younger children compared with high schoolers and adults. Third, these age-related differences would reflect not only differences in amount, but also in kind of information packaging. Fourth, our general model for characterizing developmental and cross-modal distinctions was expected to apply similarly to texts in both English and Hebrew.

Findings are presented for 52 out of a total 320 English and Hebrew narrative texts. Each text was analyzed in terms of informative (eventive, descriptive, interpretive) versus non-informative (non-novel contentive, non-contentive, extra-contentive) elements, with non-contentive material broken down into disfluencies and four types of discourse markers. Tables 1 to 3 sum up findings for the spoken and written narrative texts of three groups of English- and Hebrew-speaking schoolchildren (8 in grade school, 8 junior high, and 8 high school [48 texts], plus two adults [4 texts]).

Table 1 presents raw figures for average text length in terms of number of clauses per text across the variables of age, modality, and language.

*Table 1.* Mean number of clauses in spoken and written texts in English and Hebrew, by age group

Group	Modality # Texts	Spoken		Written	
		English	Hebrew	English	Hebrew
G [9-10]	8 + 8	114	93	67	59
J [12-13]	8 + 8	70	84	79	65
H [16-17]	8 + 8	179	113	176	68
Adults	2 + 2	33	28	25	21

Table 1 shows that, except in one case, English texts are longer than the Hebrew, a trend that was strongly confirmed by subsequent statistical analysis of a total 160 written and spoken narrative texts in the two languages (see endnote 6). Because of the unequal length of texts, informative versus noninformative material that they contained was analyzed by proportional distributions – as shown in the following tables.

Tables 2a and 2b present the mean percentage of informative versus non-informative units by age group, modality, and language, out of the total number of units counted in these 52 texts.

*Table 2a.* Distribution of informative material in spoken and written English and Hebrew texts (N=52), by age group

Modality Group	Spoken			Written		
	English	Hebrew	Average	English	Hebrew	Average
G [9-10]	57.6	55.6	56.6	94.6	84.8	89.7
J [12-13]	50.4	54.7	52.5	88.5	88.7	88.6
H [16-17]	47.3	48.2	47.8	82.8	80	81.4
Adults	52.9	46	49.5	85	76.7	80.8

*Table 2b.* Distribution of non-informative material in spoken and written English and Hebrew texts (N=52), by age group

Modality Group	Spoken			Written		
	English	Hebrew	Average	English	Hebrew	Average
G [9-10]	42.4	44.4	43.4	5.4	15.2	10.3
J [12-13]	49.6	45.3	47.5	11.4	11.3	11.3
H [16-17]	52.7	51.8	52.2	17.2	20	18.6
Adults	47.1	54	50.5	15	23	19

The figures in Tables 2a and 2b confirm our predictions regarding *modality*: Written texts indeed reveal higher informativeness and hence denser information packaging than their spoken counterparts as measured here: 80–90% of written text material consists of informative units, compared with only around half of the spoken text material. As predicted, this distribution is stable across English and Hebrew, as averaged for each modality. Tables 2a and 2b further reveal a slight tendency towards an increase in non-informative material in the older age groups in both speech and writing. (Subsequent statistical analysis of the larger sample reveals that this trend exists, but is non-significant). Relatedly, there appeared to be

a difference in the *quality* of non-informative units as a function of age, as shown below.

Table 3 presents the mean percentage of each type of informative units out of the total number of informative text units, by age group, modality, and language.

*Table 3.* Distribution in percentages of types of informative material (eventive, descriptive, interpretive) in spoken and written English and Hebrew texts (N=52), by age group

Modality	Spoken				Written		
	Group	ENG	HEB	Average	ENG	HEB	Average
Eventives	G	45.8	56	50.9	53.9	48.7	51.3
	J	36.7	34.6	35.6	35.2	36.2	35.7
	H	40.4	16.2	28.3	34.2	21.2	27.7
	Adults	33.3	34.8	34	35.3	39.1	37.2
Descriptives	G	51.4	40	45.7	42.3	46.2	44.3
	J	60	48.1	54	57.4	55.3	56.4
	H	48.1	61.8	54.9	47.7	59.6	53.7
	Adults	44.4	34.8	39.6	41.2	34.8	38
Interpretives	G	2.8	4	3.4	3.9	5.1	4.5
	J	3.3	17.3	10.3	7.4	8.5	8
	H	11.5	22.1	16.8	18	19.2	18.6
	Adults	22.2	21.7	21.9	23.5	26.1	24.8

The figures in Table 3 bear out our predictions. The distribution of Infus is similar across English and Hebrew, especially in the written texts. Further, Table 3 underscores the *developmental* aspect of our analysis: The averaged means across the two languages reveal that, with increased age and literacy, the amount of eventives goes down, with a concomitant rise in descriptives and interpretives (supported by statistics from the larger sample). This change is especially marked in the interpretive component, which increases dramatically from under 5% to close to one quarter of the informative material, particularly in the written texts.

Our last analysis is of the three types of non-informative material by modality, age group, and language. Table 4 shows their frequencies in raw scores, in terms of text length measured by number of clauses, since there were relatively few such elements, particularly in the written texts.

Table 4. Distribution of types of Non-informative material in spoken and written English and Hebrew texts (N=52), by age group

Modality Non-Informative Units: Frequency	Spoken		Written		
	Group	ENG	HEB	ENG	HEB
(1) Non-novel contentives: <i>Reiterations</i>	G	12	5	1	2
	J	4	6	1	1
	H	10	13	2	2
	Adults	2	6	0	2
(2) Non-contentives: <i>Disfluencies</i>	G	20	16	0	0
	J	13	14	1	1
	H	22	15	1	0
	Adults	3	5	0	0
(2) Non-contentives: <i>Discourse Markers</i>	G	21	19	1	2
	J	40	23	3	4
	H	78	40	11	9
	Adults	10	16	3	5
(3) Extra-contentives: <i>Story-Externals</i>	G	0	0	1	3
	J	2	1	2	0
	H	6	5	9	2
	Adults	1	0	0	0

The distributions in Table 4 reveal clear modality effects: Spoken texts contain many reiterations, disfluencies, and discourse markers, while written texts contain almost no collateral material, and far fewer discourse markers. Adolescents (group J and especially H) use relatively more such elements than the two other groups. This apparent U-shaped curve in use of discourse markers in both spoken and written texts is strongly confirmed by findings from the larger sample, which included 10 subjects in all four age groups, in both languages. And again, the tendency for extra-contentive, expository-like reflective comments to increase with age and literacy, is confirmed when more adult texts are added.

#### 4. Discussion

This analysis has focused on modality-driven differences between spoken and written narratives produced by speaker-writers of American English and Israel Hebrew. In our proposed model of narrative informativeness, Eventive information units form the story core, anchoring Descriptive and

Interpretive novel contentive units as satellites within the story frame. Across and beyond this frame, non-contentive and extra-contentive units governed by factors of ongoing text production express listener-reader orientation and inter-textual commentary.

Our preliminary findings indicate that this model effectively captures essential differences between spoken and written narratives across development: Spoken texts contain twice, sometimes four times, as many “non-informative” elements as their written counterparts. These non-informative elements are mostly either non-novel reiterations or non-contentive “collateral” type material, peaking in adolescence. With age, speaker-writers produce information units that are less eventive and more evaluative in nature, with story-external generalizations and (inter)textual commentary constituting a larger part of “non-informative” material.

In terms of modality, spoken texts are across age groups far more cluttered and less informative than their written counterparts, in the sense in which the term “informative” was defined in the present context. Within the story-frame, spoken texts contain more reiterations of contentive material, while across the story frame, they contain more communicatively- and processing-motivated material of the kind we termed “non-novel” or “non-contentive”.

Note that our prediction that with age and increased literacy we would find less difference in level of informativeness between the written and spoken texts was *not* confirmed. We interpret the fact that quantitatively, the overall ratio of informative to non-informative material remains the same across age-groups as due to the impact of the differences in processing constraints and circumstances of online versus offline text production that apply in speech compared with writing (see Section 1). That is, these types of modality effects appear to be constrained by general cognitive factors that are in operation from early on. On the other hand, there is a marked *qualitative* change in the types of non-informative material as a function of age. With increased age and literacy, recognition of the requirements of written language as a special style of discourse makes narrators increasingly constrained in use of highly colloquial type discourse markers (interactive, segment tagging, and qualifying) as they come to adopt an increasing text-oriented perspective and a more removed discourse stance as befitting the more formal and highly monitored nature of written discourse (Berman 2005; Berman et al. 2002). Interestingly, this difference in treatment of written text construction with increasing age and literacy is found even in the genre of personal experience narratives, known to elicit

less formal and high-level language usage than expository type prose (Bar-Ilan and Berman in press; Berman and Nir-Sagiv 2004).

Moreover, these differences cut across the two languages, suggesting that in western cultures, at all events, differences in basic processing of the two modalities applies in very similar ways. This finding is consistent with comparisons of written and oral narrative productions of schoolchildren and adolescents in Swedish (Strömquist, Wengelin, and Nordqvist 2004) and French (Gayraud 2000). On the other hand, there is some reason to expect that written language as a special style of discourse rather than as a special mode of production might be more affected by cross-cultural differences. For example, French schooling requires and inculcates very different styles of expression in writing compared with speech (Jisa 2004), compared with the more egalitarian school systems in, say, Sweden and the United States, whereas official written Hebrew differs markedly from colloquial spoken usage but this is not directly taught in the schools. These issues were not addressed in the present study and are worthy of further consideration.

Analysis revealed a clear age-related change shared across English and Hebrew in the nature (although, again, not in the proportion) of contentive narrative material in our data-base. Development reveals a steady elaborating of bare-bones narrative information: Reference to descriptive and, among older narrators, interpretive elements increases to flesh out the eventive narrative frame, embedding it in the attendant circumstances and motivations that form the background to events. And more mature, literate narrators enrich their accounts by “story-external” commentary, evidence of an increased ability to adopt a more distanced, text-oriented discourse stance. These developments are driven by an internalized narrative schema and developing storytelling abilities that are common to speakers of different languages, illustrated here by Californian English and Israeli Hebrew.

These findings support the idea that discourse development involves greater variety of textual information, in the sense of divergence from canonic, genre-typical elements of text construction – recounting of past events in narratives or formulating of timeless generalizations in expository discussions. The narratives of grade school children start out with a single, mainly eventive, informational layer, which is gradually expanded and elaborated by adding layers of descriptive and interpretive content.

## References

- Aisenman, Ravid, and Nurit Assayag  
1999 Fact and fiction in evaluating narrative evaluation. In *Working Papers in developing literacy, Volume 1*, Ravid Aisenman (ed.), 38-57. Tel Aviv: Tel Aviv University.
- Bamberg, Michael, and Robin Damrad-Frye  
1991 On the ability to provide evaluative comments: Further explorations of children's narrative competence. *Journal of Child Language* 18: 689-710.
- Bar-Ilan, Laly, and Ruth A. Berman  
In press Developing register differentiation: The Latinate-Germanic divide in English. *Linguistics*.
- Berman, Ruth A.  
1997 Narrative theory and narrative development: The Labovian impact. *Journal of Narrative and Life History* 7: 235-244.  
1998 Typological perspectives on connectivity. In *Issues in the theory of language acquisition*, Norbert Dittmar and Zvi Penner (eds.), 203-224. Bern: Lang.  
2001 Setting the narrative scene: How children begin to tell a story. In *Children's language Volume 10*, Ayhan Aksu-Koç, Caroline Johnson, and Keith Nelson (eds.), 1-31. Mahwah, NJ: Lawrence Erlbaum.  
2005 Introduction: Developing discourse stance in different text types and languages. In *Journal of Pragmatics* 37 (2): 105-124 (special issue on *Developing Discourse Stance across Adolescence*).
- Berman, Ruth A., Bracha Nir-Sagiv, and Laly Bar-Ilan  
In press Vocabulary development across adolescence: Text-based analyses. In *Language education in Israel: Papers in Honor of Elite Olshtain*, Irit Kupferberg and Anat Stavans (eds.). Jerusalem: Magnes Press.
- Berman, Ruth A., Hrafnhildur Ragnarsdóttir, and Sven Strömqvist  
2002 Discourse stance. *Written Languages and Literacy* 5: 255-290.
- Berman, Ruth A., and Dan I. Slobin  
1994 *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Berman, Ruth A., and Ludo Verhoeven  
2002 Crosslinguistic perspectives on developing text production abilities. *Written Language and Literacy* 5: 1-44.

- Biber, Douglas, Susan Conrad, and Randi Reppen  
1998 *Corpus linguistics: Investigating language and use*. Cambridge: Cambridge University Press.
- Brinton, Laurel  
1996 *Pragmatic markers in English: Grammaticalization and discourse function*. New York: Mouton de Gruyter.
- Chafe, Wallace  
1994 *Discourse, consciousness, and time: The flow of language in speech and writing*. Chicago: Chicago University Press.
- Clark, Herbert H.  
1996 *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert H., and Jean E. Fox Tree  
2002 Using uh and um in spontaneous speaking. *Cognition* 84: 73–111.
- Clark, Herbert H., and Richard J. Gerrig  
1990 Quotations as demonstrations. *Language* 66: 764-805.
- Coulmas, Florian  
1986 Reported speech: some issues. In *Direct and Indirect Speech*, Florian Coulmas (ed.), 1-28. Berlin: Mouton de Gruyter.
- Fernandez, Josie  
1994 *Les particules énonciatives*. Paris: Presses Universitaires de France.
- Fleischman, Suzane, and Marina Yaguello  
2003 Discourse markers across languages: Evidence from English and French. *Discourse across languages and cultures*, Carol Lynn Moder and Aida Martinovic-Zic (eds.), 130-147. Amsterdam: John Benjamins.
- Gayraud, Frédérique  
2000 Le développement de la différenciation oral/écrit vu à travers le lexique. Ph.D. diss., Université Lumière Lyon 2.
- Haiman, John, and Sandra A. Thompson (eds.)  
1988 *Clause combining in grammar and discourse*. Amsterdam: John Benjamins.
- Halliday, M.A.K.  
1989 *Spoken and written language*. Oxford: Oxford University Press.
- Hopper, Paul  
1979 Aspect and foregrounding in discourse. *Syntax and semantics, Volume 12*, Talmy Givón (ed.), 213–242. New York: Academic Press.

- Jisa, Harriet  
2004 Growing into academic French. In *Language development across childhood and adolescence*, Ruth A. Berman (ed.), 135-162. Amsterdam: John Benjamins.
- Jucker, A.H., and Yael Ziv (eds.)  
1998 *Discourse markers: Description and theory*. Amsterdam: John Benjamins.
- Labov, William  
1972 *Language in the inner city*. Philadelphia: University of Pennsylvania Press.  
1997 Some further steps in narrative history. *Journal of Narrative and Life History* 7: 394-415.
- Lambrecht, Knud  
1994 *Information structure and sentence form: A theory of topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- Malvern, David D., Brian J. Richards, Ngoni Chipere, and Pilar Durán  
2004 *Lexical diversity and language development: Quantification and assessment*. London: Palgrave Macmillan.
- Pacton, Sébastien, and Michel Fayol  
2004 Learning to spell in a deep orthography: The case of French. In *Language development across childhood and adolescence*, Ruth A. Berman (ed.), 163-176. Amsterdam: John Benjamins.
- Prince, Ellen  
1981 Towards a taxonomy of given-new information. *Radical pragmatics*, Peter Cole (ed.), 223-255. New York: Academic Press.
- Ravid, Dorit  
2004 Emergence of linguistic complexity in written expository texts: Evidence from later language acquisition". In *Perspectives on language and language development*, Dorit Ravid and Hava Bat-Zeev Shyldkrot (eds.). Dordrecht: Kluwer.
- Ravid, Dorit, and Ruth A. Berman  
In press Information density in the development of spoken and written narratives in English and Hebrew. In *Discourse Processes* 41 (2) (special issue on *Surface cues of content and tenor in texts*).
- Ravid, Dorit, and Liliana Tolchinsky  
2002 Developing linguistic literacy: A comprehensive model. *Journal of Child Language* 29: 419-448.

- Reinhart, Tanya  
1984 Principles of gestalt perception in the temporal organization of narrative texts. *Linguistics* 22: 779–809.  
1995 From text to meaning: Evaluative devices. In *Cognitive aspects of narrative structure*, Yeshayahu Shen (ed.), 4-37. Tel Aviv University: Porter Institute. [in Hebrew]
- Schiffirin, Debra  
1995 *Discourse markers*. Cambridge: Cambridge University Press.
- Segal, Michal  
2001 Form-Function relations in the expression of evaluation in personal-experience narratives across adolescence. Ph.D. diss., Tel Aviv University. [in Hebrew]
- Strömqvist, Sven, Åsa Nordqvist, and Åsa Wengelin  
2004 Writing the frog story: Developmental and cross-modal perspectives. In *Relating events in narrative: Typological contextual perspectives*, Sven Strömqvist and Ludo Verhoeven (eds.), 359-394. Mahwah, NJ: Lawrence Erlbaum.
- Tolchinsky, Liliana  
2003 *The cradle of culture: What children know about writing and numbers before being taught*. Mahwah, NJ: Lawrence Erlbaum.
- Vallduví, Enric, and Elisabeth Engdahl  
1996 The linguistic realization of information packaging. *Linguistics* 33: 459–519
- Wengelin, Åsa, and Sven Strömqvist  
2004 Text-writing development viewed through on-line pausing in Swedish. In *Language development across childhood and adolescence*, Ruth A. Berman (ed.), 177-190. Amsterdam: John Benjamins.

## Notes

---

- \* This paper is a revised and expanded version of a talk given at The Georgetown University Round Table [GURT2003] on Language in Use: Discourse and Cognitive Perspectives to Language and Language Learning, Washington D.C., February 2003. This talk formed the basis of subsequent investigation by the authors of the present study, adopting a rather different focus than the present chapter (Ravid and Berman, in press). The authors are indebted to Bracha Nir-Sagiv for invaluable help in coding and analysis of the data and in preparing the revised manuscript, which benefited greatly from editorial input from Dr. Andrea Tyler and the insightful comments of an anonymous reviewer. Inadequacies that remain are ours alone.
1. Our data thus combine features of both what are termed “natural” and “contrived” data in a recent debate in the journal *Discourse Studies* 4, 4 (2002), pp. 511-548.
  2. In examples, punctuation and orthography are standardized from the original transcripts.
  3. Hebrew texts are given in free translation rather than morpheme-by-morpheme glosses. For example, progressive and perfect aspect is indicated, although Hebrew does not have grammatical aspect, and so is the indefinite article. But we try to retain some flavor of the original register of usage and rhetorical text even where this results in awkward English.
  4. Our analysis excludes nonlexical filler syllables like English *er, um* (Clark and Fox Tree 2002).
  5. The project on Developing Literacy in Different Languages and Different Contexts was supported by a major grant from the Spencer Foundation, Chicago, to Ruth Berman as PI. Data collection was supervised by Dorit Ravid for Hebrew in Israel and by Judy S. Reilly for in San Diego, California.
  6. Since submitting this paper for publication, the authors have concluded investigation along largely similar lines of 160 texts in English and Hebrew. Statistically-based analyses reported in Ravid and Berman (in press) fully substantiate the predictions and initial findings presented here.