

# On newspaper headlines as relevance optimizers

Daniel Dor\*

*Department of Communications, Tel Aviv University, Tel Aviv, Israel*

---

## Abstract

This paper suggests an explanatory functional characterization of *newspaper headlines*. Couched within Sperber and Wilson's (1986) *relevance* theory, the paper makes the claim that headlines are designed to *optimize the relevance* of their stories for their readers: Headlines provide the readers with the optimal ratio between contextual effect and processing effort, and direct readers to construct the optimal context for interpretation. The paper presents the results of an empirical study conducted in the news-desk of one daily newspaper. It shows that the set of intuitive professional imperatives, shared by news-editors and copy-editors, which dictates the choice of headlines for specific stories, can naturally be reduced to the notion of relevance optimization. The analysis explains why the construction of a successful headline requires an understanding of the readers—their state-of-knowledge, their beliefs and expectations and their cognitive styles—no less than it requires an understanding of the story. It also explains the fact that skilled newspaper readers spend most of their reading time *scanning* the headlines—rather than reading the stories.

© 2002 Elsevier Science B.V. All rights reserved.

*Keywords:* Headlines; Relevance theory; Pragmatics; News value; News framing; Media, communication

---

## 1. Introduction

This paper is an attempt to suggest an explicit and generalized answer to a very fundamental question in the study of the mass media, i.e., the question of the *communicative function of newspaper headlines*. The importance of the role of headlines in the communicative act performed by newspapers can hardly be exaggerated, yet the nature of this role has virtually never been explicated in the literature. As we shall see below, the regular strategy adopted in the literature has been to make fine-grained *descriptive* distinctions between different *types* of headlines—news headlines in 'quality newspapers'; news headlines in 'tabloid newspapers'; 'summarizing

---

\* Tel.: +972-3-6406521; fax: +972-3-6406032.

*E-mail address:* danieldd@post.tau.ac.il (D. Dor).

headlines'; 'localizing headlines', 'quotation headlines', etc.—and assign them different types of communicative functions. In this paper, I will suggest an explanatory functional definition of newspaper headlines which attempts to transcend the above distinctions in type and explain the very fact that newspapers—all types of newspapers—have headlines in them. The functional definition to be developed in this paper relies very heavily on Sperber and Wilson's (1986) technical notion of *relevance*. Newspaper headlines will be functionally defined as *relevance optimizers*:

*Newspaper headlines are relevance optimizers: They are designed to optimize the relevance of their stories for their readers.*

This functional definition positions the headline in its appropriate role as a textual *negotiator* between the story and its readers. It explains why the construction of a successful headline requires an understanding of the readers—their state-of-knowledge, their beliefs and expectations and their cognitive styles—no less than it requires an understanding of the story. It reduces the differences between the different subtypes of headlines mentioned above to a matter of tactical choice: As we shall see, all the different subtypes target the same functional *goal*, that of relevance optimization, although they do it in different ways.

The literature on newspaper headlines covers a wide range of theoretical and empirical topics, all the way from the grammar of English headlines to the effects of headlines on news comprehension and recall.<sup>1</sup> Surprisingly, however, the literature dealing directly with the communicative function of headlines is rather sparse. I will review it in the next section. In Section 3, I will briefly introduce Sperber and Wilson's theory, and then develop the notion of *relevance optimization*. In Section 4, I will apply the notion of relevance optimization to newspaper headlines. In Section 5, I will present the results of an empirical study conducted in the news-desk of the Israeli national newspaper *Ma'ariv*, where I followed the process of headline production from close range.<sup>2</sup> I will show that the set of intuitive professional imperatives, shared by news-editors and copy-editors, which dictates the choice of headlines for specific stories, can naturally be reduced to one meta-imperative: *Make the headline such that it renders the story optimally-relevant for the readers*. In Section 6, I will apply the relevance-based conception to the analysis of tabloid headlines. In Section 7, I will deal with the role of the reader in this framework, and show that my relevance-based theory explains some of the more intriguing behavioral patterns manifested by newspaper readers—especially the fact that many skilled readers

<sup>1</sup> On headline reading, interpretation and recall, see Henley et al. (1995), Leon (1997), Lindemann (1989), Perfetti et al. (1987), Pfau (1995) and van Dijk (1988 and references therein); on headline production, see Bell (1984, 1991), Fasold (1987) and Chang et al. (1992); on the grammar of headlines, see Bell (1984), Jenkins (1990) and Mardh (1980); on metaphors in headlines, see de Knop (1985); on headlines from a cross-linguistic perspective, see Dierick (1987) and Sidiropoulou (1995).

<sup>2</sup> Between 1996 and 1998, I worked as a senior news-editor and head of the news-desk in *Ma'ariv*. This was a period of very intensive *participant observation*: I was involved in the decision-making process concerning the formulation of thousands of headlines. The e-mail exchanges which were analyzed for this paper were randomly collected throughout this period—from other senior editors.

spend most of their reading time *scanning* the headlines rather than reading the stories. In the concluding section, I will sketch some of the larger-scale implications of my theory, and suggest some directions for further research.

## 2. Multiple types, multiple functions

Traditionally, newspaper headlines have been functionally characterized as short, telegram-like *summaries* of their news items. This is especially true with respect to *news* headlines. Van Dijk (1988) couches this traditional insight within his discourse-analytic framework: “Each news item in the press has a Headline . . . and many have a Lead, whether marked off by special printing type or not. We also have an elementary rule for them: Headline precedes Lead, and together they precede the rest of the news item. Their structural function is also clear: Together they express the major topics of the text. That is, they function as an initial summary. Hence, as in natural stories, we may also introduce the category Summary, dominating Headline and Lead. The semantic constraint is obvious: Headline + Lead summarize the news text and express the semantic macrostructure.”

Obviously, some newspaper headlines do provide what seems to be a summary (or abstract) of their stories, but the general theoretical conception which takes this to be the essential function of the headline seems to be too narrow, for at least three complementary reasons. First, even the most prototypical news headlines, those which appear in what is sometimes called ‘quality newspapers’, do not always summarize their stories. Some headlines *highlight* a single detail extracted out of the story, and others contain a *quotation* which the editor decided should be promoted to the foreground. As we shall see below, some headlines even contain material which does *not* appear in the news item itself. The fact that headlines do not always summarize, but sometimes highlight or quote, has been noted by different writers. Bell (1991), for example, makes a distinction between headlines which “abstract the main event of the story”, and headlines which “focus on a secondary event or a detail” (p. 188–9). Nir (1993) distinguishes between headlines which function as “a summary of the story” and “headlines which, rather than summarize the story, promote one of the details of the story” (p. 25).<sup>3</sup>

Second, the traditional notion of headlines-as-summaries definitely does *not* capture the function of headlines in more popular newspapers, and especially in tabloids. This point has been made by different writers, most notably by Lindemann (1990). As Lindemann shows, tabloid headlines rarely summarize their stories, are not always telegram-like, and in many cases are not even informative. Lindemann discusses the

<sup>3</sup> Note that none of the above writers goes beyond the descriptive labeling of the different types of headlines to suggest explicit theoretical definitions and explanations. This fact is most clearly demonstrated by Bernstein and Garst (1982), quoted in Lindemann (1990), who claim that “... the headline contains the main highlight of the story. Since it is the most conspicuous part and the part that is read first, the copy editor must present the essence of the news before he goes further”. In this short quotation, Bernstein and Garst seem to equate the *essence* of the story with its *highlight*, thus equating the function of summarizing with that of highlighting.

function of tabloid headlines in *poetic* terms: They present the reader with a “fairly complex riddle”, which, first, triggers frames and belief systems in the reader’s mind, and, then, gets resolved in the ensuing text. Thus, the following headline,

(1) NO-LA-LA! The Frogs Get Bored with Bed

traps the reader “in the treadmill of well-established cliches and prejudice”, through the use of such expressions as *frogs*, *no-la-la* and *bed*, and is then informationally resolved in the intro: “The days of the great French lovers are over—froggies just don’t fancy it any more. A third of women and a quarter of men told a nationwide survey they found bedtime one big yawn”.

Implicit in Lindemann’s analysis is the assumption, that the function of tabloid headlines is so radically different from their function in quality newspapers, that the two cannot be theoretically unified. As I will show below, the relevance-based analysis will allow exactly for that—to my mind, a very welcome theoretical result.

The third reason to reject the traditional conception is the simple fact that headlines seem to have an additional, *pragmatic* function, beyond the semantically-oriented function which is supposed to be captured by the headline-as-summary analysis. Bell (1991) says that headlines are a “part of news rhetoric whose function is to attract the reader” (p. 189). Nir (1993) claims that the headline has “to attract the attention of the reader and provoke the reader to read the whole story”. In a sophisticated analysis of the semiotics of headlines, Iarovici and Amel (1989) explicitly contend that the headline has a “double function”:

“The implicit convention between author and reader regarding the intention of correlating a text to another text as a headline, and regarding the formal marking of this quality by a privileged position, concerns the double function of the headline: a *semantic function*, regarding the referential text, and a *pragmatic function*, regarding the reader (the receiver) to whom the text is addressed. The two functions are simultaneous, the semantic function being included in and justified by the pragmatic function. ... The main function of the headline ... is to alert the reader (receiver) to the nature or the content of the text. This is the pragmatic function of the headline, and it includes the semantic one. The headline enables the reader to grasp the meaning of the text. The headline functions as a plurality of speech acts (urging, warning, and informing)” (p. 441–443).

The challenge posed by the above assertions is that of *theoretical unification*. At least two questions are involved: First, can we functionally define the headline in a way which would transcend the above distinctions between the different *semantically-oriented* functions? In other words, is there a generalized function which summarizing headlines, localizing headlines and quotation headlines have in common? Second, can we define the headline in a way which would transcend the distinction between the above semantic function and the parallel *pragmatic* function which

headlines fulfill? I would like to claim that this theoretical move becomes possible once we couch the functional analysis of headlines within the framework of Sperber and Wilson's (1986) theory of *relevance*.<sup>4</sup>

### 3. Relevance theory and relevance optimization

Sperber and Wilson's theory is an attempt to reduce a very complex set of phenomena having to do with communication and interpretation to a very constrained set of explanatory, *cognitive* notions. In its essence, the theory is one of cognitive *cost-effectiveness*: It claims that human cognitive processes are geared to achieving the greatest possible cognitive effect for the smallest processing effort. This meta-principle is incarnated in Sperber and Wilson's technical notion of *relevance*. Let us take a look at the fundamental tenets of this framework:

Our starting point is the individual *mind*: Every individual mentally represents in his or her mind a huge set of *assumptions*. Assumptions are propositional entities—they are the type of entities that can be believed to be true. Our assumptions may include, among other things, information on the immediate physical environment, expectations about the future, scientific hypotheses, religious beliefs, anecdotal memories, general cultural assumptions, beliefs about the personal lives of our acquaintances, knowledge about politics and history, beliefs about our own emotions, fears and hopes, and so on. Each of the assumptions represented by the individual has a “strength” for that individual. The strength of the assumption for the individual is the level of confidence with which the individual holds to the belief that the assumption is true. The strength of the assumption is a function of its cognitive *processing history*. Thus, for example, “assumptions based on a clear perceptual experience tend to be very strong; assumptions based on the acceptance of somebody's word have a strength commensurate with one's confidence in the speaker; the strength of assumptions arrived at by deduction depends on the strength of the premises from which they were derived” (p. 77) Note that the strength of an assumption for the individual has nothing to do with its *objective* validity—individuals may have a very strong belief in assumptions which are totally false, and vice versa.

When an individual hears, or reads, a novel assumption, he or she always interprets it in a *context*. The notion of context is used here as a *psychological construct*: It is a subset of the assumptions which the hearer already represents in his or her long-term memory. Informally, what the mind of the individual does in the process of interpretation may be thought of as a *comparison* of the new assumption with the subset of assumptions represented in the individual's memory. Sperber and Wilson name the cognitive apparatus responsible for this process of comparison—“the deductive device”. The comparison of the novel assumption with the existing

---

<sup>4</sup> The general notion of pragmatic relevance, which is not to be equated with Sperber and Wilson's technical one, plays some role in van Dijk's (1988) analysis of news *selection*. However, van Dijk does not make the connection between his notion of relevance and the function of headlines, which he takes to be summaries of their texts.

assumptions may have different types of *outputs*: It may turn out, for example, that the novel assumption already exists in the individual's long term-memory, in which case it is not *new* for the individual. Or it may be new, in which case it may either be in line, or in contradiction with some of the already existing assumptions. If, for example, the novel assumption contradicts existing assumptions, and if it is strong enough, the process of comparison will end up with the *weakening* of the existing assumptions. In some cases, it may even end up with the *erasure* of those assumptions. If the new information is in line with some existing assumptions, it may serve to *strengthen* them some more. Moreover, the union of the new assumption with some existing assumptions may lead to the deduction of *additional* assumptions. Thus, for example, if the individual already represents the assumption that “whenever Peter goes to a party, it becomes a success”, and he or she now learns that “Peter came to Bill's party”, then the deductive device deduces an additional assumption, namely that “Bill's party was a success”. To the extent that the comparison of the new assumption with the old ones results in a *change* to the individual's set of prior assumptions (if it either adds new assumptions, or weakens or strengthens existing ones), we say that the new information has a *contextual effect* for the individual.

Now, the following point is crucial: The deductive device does *not* compare every novel assumption to the *entire* set of assumptions represented in the individual's long-term memory. Doing this would be cognitively impossible. This means that the comparison is done with some *subset* of existing assumptions. This, in turn, raises a very important question: How does the deductive device choose this subset? Traditionally, pragmaticists have assumed that the context for the interpretation of an utterance is simply *given*: It consists of the immediate environment and the information explicitly mentioned in the conversation prior to the utterance. Sperber and Wilson flip this assumption on its head and suggest a radical alternative: They show that the deductive device has to *update* the context for the interpretation for each new assumption, and that the specific subset of existing assumptions which is chosen for the context is *determined*, at least partially, by the content of the new assumption. In cognitive terms, this means that the order of events in comprehension is reversed: It is not that the deductive device first sets the context, and then interprets the new assumption. On the contrary, the deductive device has to partially figure out the meaning of the new assumption, retrieve a specific subset of assumptions from long-term memory, store them in its own short-term memory, and then make the comparison. An example should make this radical conception rather intuitive. Take a look at the following exchanges:

(2) A: How are you?

B: Not so good, Mary has that ear-infection again, I'm worried.

(3) A: How are you?

B: Great, I just bought the tickets. We're flying to Beijing in exactly four weeks.

In order to interpret B's answer in each of these exchanges, A has to compare them to a subset of existing assumptions. The proper context in (2) should include assumptions about the identity of Mary, her relation to B, her medical history, ear-

infections, and so on and so forth. The proper context in (3) should include assumptions about B's travel plans, the identity of her companion, or companions, whatever assumptions A has about Beijing, and so on and so forth. Obviously, these assumptions are *not* stored in A's short-term memory on a permanent basis. A's deductive device has to retrieve these assumptions from long-term memory, and only then make the comparison and deduce the contextual effects.

We may now make two parallel cognitive assumptions regarding the process I have described. First, we may assume that in its appropriate context, a new piece of information has a certain *number* of contextual effects, which, at least theoretically, can be counted. Practically speaking, when we deal with interpretations of actual utterances by real people, we do not know exactly how to make the measurement, but the idea itself is intuitive enough for us to accept. We may be pretty certain that in different contexts, the same piece of information may yield different amounts of contextual effects, and that in the same context, some pieces of information would yield more contextual effects than others.

Second, we may assume that the work of the deductive device involves some mental *effort*, which—theoretically speaking, again—may be measured.<sup>5</sup> Other things being equal, for example, the computation of a more complex piece of information will take more effort than the computation of a simpler one. Moreover, the construction of a new context for interpretation also involves some mental effort: To the extent that the interpretation of the novel piece of information necessitates the retrieval of a larger set of assumptions from long-term memory, the mental effort involved in the interpretation process would be greater.

The measurements of contextual effect and mental effort constitute the basis of Sperber and Wilson's notion of *relevance*:

- (4) Relevance for an individual (p. 145):
- a. An assumption is relevant to an individual to the extent that the contextual effects achieved when it is optimally processed are large.
  - b. An assumption is relevant to an individual to the extent that the effort required to process it optimally is small.

It is crucial to understand that this is not a definition of relevance in some objective sense, but a claim concerning the way our minds make relevance *judgments* about new assumptions: We consider new assumptions to be relevant if they carry a contextual effect at a reasonable cognitive price. We judge new assumptions to be irrelevant if they do not carry a contextual effect, or if the computation of the contextual effect entails too much of a mental effort. Note that this is a comparative, gradual conception of relevance, rather than a binary one: New assumptions are not either relevant or not; they are more or less relevant than others, in different contexts, for different people. Thus, for example, a regular newspaper reader will prob-

---

<sup>5</sup> Sperber and Wilson (1986, p. 130) conceptualize about the measurement of contextual effects and cognitive effort in terms of *physico-chemical changes* which occur in the brain as a result of the processing of the contextual effects.

ably judge a piece of *local* news to be more relevant than a piece of foreign news because (i) the potential contextual effect derivable from the local news would probably be larger; and (ii) the effort needed to interpret the foreign news (especially in terms of the retrieval of the appropriate set of assumptions from long-term memory) would probably be larger. This judgment need not be made consciously; the reader may simply skip the foreign-news page, or note that ‘foreign news is boring’. Note, however, that the very same reader *may* take the trouble to read the foreign news to the extent that their contextual effects would be worth his or her processing effort. This may be the case, for example, if the story is about a country which the reader intends to visit; if some people which the reader knows are there; if there is a local angle to the foreign story; if the foreign story has a global consequence which is felt locally, and so on.

Finally, note that this technical notion of relevance should *not* be equated with relevance in the *ordinary* sense of the word. Relevance in this ordinary sense may be thought of as the measurement of the association, or congruence, between some content and its context of interpretation. Thus, a news story will be relevant in this sense to the extent that it is *about* those issues which are directly related to the readers’ lives and interests. Indeed, relevance in this sense may play a role in *news value* judgments. Note, however, that a story may be relevant in this ordinary sense but very low on relevance in the technical sense (if it is long and complicated to read, for example, or if it does not carry a lot of new information); and it may be irrelevant in the ordinary sense, but high on relevance in the technical sense— if its potential contextual effects justify the construction of a new context for interpretation.

Now, our technical definition of relevance is addressee-oriented, but it may actually tell us something of importance about the role of *speakers* in communicative contexts. Think about a speaker, Ann, who is trying to tell her addressee, John, a story. Being a cooperative communicator, Ann would like to make the story as relevant for John as possible. How should she go about achieving this goal? According to relevance theory, she has three principled strategies which she can try to adopt:

- (i) First, Ann can try to compress the largest possible number of new assumptions (those which are new for *John*) into her story: Other things being equal, the more new assumptions the story contains, the more contextual effects it may have for John. In the worst-case scenario, the story will not contain any assumptions which are new for John, in which case he is going to find it totally irrelevant. Ann definitely needs to find a way to do better than that. In the best-case scenario, on the other hand, the story will contain a very large number of new assumptions. As we shall see below, this is not always going to be possible.
- (ii) Second, Ann can try to minimize John’s processing effort: Other things being equal, the smaller the effort he has to put in, the greater the relevance of the story is going to be for him. In the worst-case scenario, the story is going to be too long and complicated, and John is going to lose interest. Ann definitely has to avoid that. In the best-case scenario, on the other hand, the



story will take a very minimal effort to process. Again, this is not always going to be possible.

- (iii) Third, Ann can try to manipulate the context in which John is going to interpret her story. Other things being equal, the closer the context of interpretation is to the optimal one, the more contextual effects the story is going to carry for John. (Note that in regular conversation we regularly take the trouble to manipulate our addressee's context of interpretation, especially when we wish to "change the topic of conversation": We say things like 'Oh, I wanted to tell you something about Bill', or 'talking about school, did you hear about Bill's exam?'). In the worst-case scenario, John is going to try to interpret Ann's story in the wrong context, and the interpretation is going to yield no contextual effects. In the best-case scenario, the story is going to be interpreted in the most appropriate context, yielding the maximal amount of contextual effects. This, again, is not always going to be possible.

Now, it is very important to realize that the three strategies mentioned above are not only completely *intertwined*, but are also in direct *competition* with each other. This is why achieving the maximal results associated with each of the strategies is not always possible. This is so for the following reasons:

- (i) First, every new assumption which the speaker adds to the story does not only contribute to the overall number of contextual effects- it also adds to the overall *processing effort*. Thus, the new assumption adds to the overall *relevance* of the story only to the extent that it clearly adds *more* contextual effect than processing effort. To the extent that the new assumption adds more to the processing effort than to the contextual effect, it actually *reduces* the overall relevance of the story. In this case, more information results in less relevance. So, the attempt to maximize relevance simply by maximizing the amount of new information is bound to end up in failure. The speaker has to figure out the *optimal* amount of information which would not result in relevance reduction due to processing effort.
- (ii) Second, Ann may definitely try to maximally reduce John's processing effort by making her story short, simple and clear, but this reduction will *not* necessarily result in maximal relevance: This is so, because the reduction in the story's complexity characteristically reduces the number of its potential contextual effects. The reduction of processing effort will enhance the relevance of Ann's story *only* to the extent that the amount of effort saved is larger than the amount of contextual effects lost. So, again, Ann cannot simply reduce John's processing effort to the minimum. She has to figure out the *optimal* amount of effort which would not result in relevance reduction due to loss of contextual effects.
- (iii) Third, the number of contextual effects which John may deduce from Ann's story is not just a function of the sheer number of new assumptions in the story, but a function of the interaction between these new assumptions and the context of interpretation. This means that Ann should not just provide

John with the optimal number of new assumptions, but also take care to provide him with those *specific* assumptions which would yield the maximal amount of contextual effects in the appropriate context, and at the very same time direct John to construct that *specific* context. This complicates our relevance considerations to a considerable extent, because the construction of the appropriate context entails a significant amount of processing effort. Consequently, in principle, the construction of the appropriate context may eventually result in relevance *reduction* due to the increase in processing effort. Thus, the construction of a *partial* context for interpretation may sometimes be the optimal strategy.

As we have seen, Ann's role as the story-teller is going to be that of relevance *optimization*: She will need to provide John with the optimal *ratio* of contextual effect and processing effort. This, I would like to claim, is exactly the generalized communicative function which newspaper headlines are supposed to fulfill: They are designed to optimize the relevance of their stories for their readers.

#### 4. Newspaper headlines as relevance optimizers

Consider the following story, from the Israeli national newspaper *Ma'ariv*:

- (5) The bodies of John Kennedy Jr., his wife Caroline and his sister-in-law Lorraine were discovered yesterday in the ocean, at a depth of 30 meters, 10 kilometers away from Martha's Vineyard Island, where they were headed on Saturday. Senator Edward Kennedy, John's uncle, arrived at the site where the bodies were found, in order to identify them. Kennedy Jr. will be buried in NY in the coming days.

This news item requires a certain amount of mental effort to interpret. To begin with, the paragraph requires some effort to read: It consists of about 70 words, and is grammatically fairly complex. Moreover, the news item requires the construction of a context for interpretation- one which includes whatever the reader knows about John Kennedy Jr., his family, their disappearance two days before, the relevant geography, and probably at least something about the Kennedys' history. As we have said before, the construction of this context takes an additional effort. Let us assume, for the sake of simplicity, that the interpretation of the entire story will require the ordinary reader to invest a certain amount of effort, let us dub it E(story). Now, to the extent that the reader manages to construct the appropriate context and read the passage, the story carries a certain amount of contextual effects: It changes a lot of factual assumptions the reader represented in his or her long-term memory (e.g., the assumption 'John Kennedy Jr. is alive and well' is replaced by 'John Kennedy Jr. died in an airplane accident'), and it changes, weakens or strengthens a great many related assumptions having to do with, for example, the inescapable tragedies of the Kennedy family, the life-styles of the rich and

famous, the blindness of fate, the risks involved in flying your own plane, and so on and so forth. Obviously, different readers will probably deduce different sets of contextual effects from the story, but for the sake of simplicity, let us assume that the ordinary reader will deduce a certain amount of contextual effects, let us dub it  $C(\text{story})$ . The relevance of the story for the ordinary reader will thus be:  $R(\text{story}) = C(\text{story})/E(\text{story})$ .

Now, let us take a look at the headline the newspaper gave to the story:

(6) John Kennedy Jr.'s body found

How much effort does the reader have to invest in interpreting the headline? Obviously, much less than  $E(\text{story})$ : The headline is a single, short and simple sentence, comprising five words, and the effort needed to read it is insubstantial. The effort needed to construct the context for the interpretation of the headline is also significantly smaller—the reader does *not* need to retrieve the sets of assumptions having to do with the geography of the story, with Senator Ed Kennedy, and so on. For the sake of simplicity, let us make the arbitrary assumption that  $E(\text{headline})$  equals 10% of  $E(\text{story})$ .

Now, how many contextual effects can the reader deduce from the headline? Surprisingly, when the headline is interpreted in its reduced context, a significantly *large* subset of the contextual effects of the entire story survive. Obviously, some things are missing—for example, the fact that Kennedy's wife and his sister-in-law were found too—but Kennedy's death, its significance within the tragic history of the Kennedy family, and the more general implications of the story are clear contextual effects of the headline. Let us adopt a conservative estimate: For the ordinary reader,  $C(\text{headline})$  equals 50% of  $C(\text{story})$ . As a simple calculation clearly shows, our estimates entail that the headline multiplies the relevance of the story by five (!). It saves much more on the processing effort than it loses on the contextual effects. This is exactly what a headline should do. A short and simple text, it optimizes the relevance of the story by minimizing processing effort while making sure that a sufficient amount of contextual effects are deducible within the most appropriate context possible. Just like Ann, our story-teller, the headline does not adopt an all-or-none strategy of either reducing processing effort to zero, or maximizing new information, or constructing the most appropriate context for interpretation. Rather, it attempts to optimize the ratio between processing effort and contextual effects- and thus optimally negotiate between the story and the ordinary reader.

Note that for the optimization of relevance to be successful, the right material should be chosen for the headline. Consider, for example, the following three clauses as alternative headlines for the Kennedy story:

- (7) a. Caroline Kennedy's body found  
 b. Sen. Edward Kennedy arrived at the crash site.

- c. The bodies of John Kennedy Jr. and his wife Caroline were discovered yesterday in the ocean, near Martha's Vineyard Island.

The first two alternative headlines (7a and 7b) are probably as easy to read as the actual one, and we may assume that they require the construction of a very similar context for interpretation. However, they do *not* carry the same amount of contextual effects as the original. The third alternative (7c) carries a slightly larger number of contextual effects than the original, but it very obviously requires *much* more processing effort. Thus, all three alternative fall short of providing optimal relevance.

Is the original headline in (6) a summarizing or a highlighting headline? It is hard to tell. The important point, however, is that from our theoretical point-of-view the summarizing-highlighting distinction is simply not that crucial: Summarizing the story is just *one* tactical approach to relevance optimization. Highlighting the most intriguing aspect of the story, or reproducing the most interesting statement quoted in the story, may have the very same result. It may turn out, for example, that the quotation or the highlighted aspect carry more contextual effects than the summary of the whole narrative. In this case, the rational thing to do would be to promote them to the headline- and thus optimize the relevance of the story for the readers. The choice between these different tactical approaches is in part a matter of the editorial style of the newspaper, and to a very large extent a matter of the experience and creativity of its editors. For every given story, some headline options are going to suggest themselves. The editor may opt for a summarizing headline, a highlighting headline or a quotation headline- depending on which type of headline will provide optimal relevance.<sup>6</sup> Moreover, the editor may manipulate the length and complexity of the headline, and its specific contents. And again, these manipulations, to a very large extent, are going to be relevance-oriented.<sup>7</sup>

In the following section, I will present the results of an empirical study conducted in the years 1996–1998 in the news-desk of the Israeli national newspaper *Ma'ariv*.<sup>8</sup> In the study, I followed the decision-making process leading to the choice of headline for a large number of news items. As the results of the study clearly indicate, the set of professional intuitions shared by the editors, concerning the properties of the 'right' or 'appropriate' headline, are theoretically reduced to our notion of rele-

<sup>6</sup> I will discuss the choice of tabloid-type headlines later on.

<sup>7</sup> An anonymous referee notes that some text manipulations may not be relevance-oriented. Thus, for example, some manipulations may have to do with spacing on the page, and others with political considerations. I agree with the first point. In Dor (2001), however, I show that relevance-oriented manipulations play an extremely important role in processes of political framing.

<sup>8</sup> There are currently three national newspapers in Israel: *Yediot Ahronot* and *Ma'ariv* are considered to be the *popular* newspapers, whereas *Ha'aretz* is considered to be the quality, high-brow one. *Yediot Ahronot* and *Ma'ariv*, however, are not tabloids in the regular, American-European sense. They contain a variety of "serious" news items which is not that different from that of *Ha'aretz*, and are distinct from it especially in writing style and graphic design. In all three newspapers, headline formulation is considered part of the editorial process, and reporters do not formulate headlines for their stories.

vance: The most appropriate headline for a news item is the one which optimizes the relevance of the story for the readers of the newspaper.

## 5. The notion of relevance and the art of headline writing

In general, news editors do not work with a very explicit definition of what headlines are, or of their communicative functions. What they do work with is a cluster of professional intuitions—gradually developed ‘in the field’, and never seriously explicated—concerning the properties of what we might call, rather informally, the ‘right’, ‘appropriate’, or ‘good’ headline. When asked to provide an explicit definition of what a headline is, senior newspaper editors usually give an answer of the type: ‘I don’t know what headlines are, but I can tell a good one when I see it’. This answer is actually a pretty accurate rendition of a very fundamental sentiment: Professional knowledge is practical, not theoretical. However, when presented with a news-item, and asked to choose a headline out of a set of alternatives, experienced news editors do so with extreme ease and efficiency. Moreover, senior editors in the *same* newspaper have a very high rate of agreement on the preferred headline. This means that experienced news editors know a great deal more about the functional properties of headlines than they ever explicate. In this sense, headline production is more similar to an artistic activity than, say, to the practice of an exact science.

This affinity with the arts is very clearly reflected in the trial-and-error process which beginning copy-editors go through as part of their on-the-job training procedure. Rather than receive their professional education in the form of explicit lecturing, beginning copy-editors in *Ma’ariv* simply start out working: They are assigned a new-item, and are asked to rewrite it and suggest a headline for it. The result is then reviewed by the senior editor in charge, who, in most cases, rejects the suggested headline and writes a different one, which eventually gets published. Sometimes, the copy-editor is asked to suggest the alternative headline, which is, again, reviewed by the editor in charge. Deadline pressure usually does not allow for long explanations: When the process is over, the copy-editor gets another story, suggests a headline, which usually gets rejected, and so on and so forth. This process goes on for years, and in a real sense never ends: In *Ma’ariv*, each and every suggested headline is sent to the senior editor in chief, in the form of an electronic message, to be approved or rejected, even if the copy-editors have years of experience behind them. Obviously, the rate of rejected headlines goes down with time, when the trained editor internalizes the set of implicit intuitions shared by the other, more experienced editors, but even very experienced editors get some of their headlines rejected some of the time: Sometimes, for example, the editor in chief knows something about the wider context of the story which the copy-editor was not aware of.

The fact that these real-time negotiations about the headlines are done in writing, by e-mail messages, allowed me to follow the process of headline formulation from very close range. I collected 134 e-mail exchanges, concerning 134 news-items, and analyzed the semantic-pragmatic differences between the rejected and approved headlines. In some cases, I asked the editor in charge to reconstruct the reasons for

the rejection of the suggested headline. I then extracted a list of ten properties, which I shall call *the properties of the appropriate headline*. I submit that this list is an accurate rendition of the set of implicit intuitions shared by experienced news editors in *Ma'ariv*. In the following section, I will present the ten properties, each with its representative example, and show that the list is actually reducible to one professional meta-imperative: *Make the headline such that it renders the story optimally-relevant*.

Three notes should be made at this point: First, the following discussion should not be thought of as an attempt to construct a theoretical framework, but as a description of a set of professional *intuitions*, shared by news editors, concerning the properties of the “appropriate headline”. In other words, I do not intend to make any significant claim concerning the theoretical status of the ten properties to be discussed below. Quite obviously, some of the properties seem to bear close resemblance to some principles discussed in the literature under the rubric of *news value* (e.g., in Galtung and Ruge, 1965; Bell, 1991); other properties may remind the reader of Grice’s conversational maxims. I will leave the elaboration of these resemblances for further research, and concentrate on the possibility of reducing the entire set of properties to the relevance-based meta-intuition mentioned above.<sup>9</sup>

Second, the properties are to be thought of as *default conditions*, rather than *obligatory* ones. It is *not* the case that *every* headline should have *all* 10 properties. It is the case that a headline which meets any of these conditions is better than a headline which does not, and a headline which meets a larger number of the conditions is better than a headline which meets a smaller number of them. Thus, for example, the first property—‘headlines should be as short as possible’—should be read as saying: ‘*other things being equal*, a shorter headline is better than a longer one’.

Finally, The headlines presented in the next section are translated from the Hebrew original. I chose to keep the translation as literally accurate as possible, and avoided translating the headlines into “headlines”, because Hebrew headlines do not usually have the telegraphic syntax characteristic of English headlines.

### 5.1. *The properties of the “appropriate headline”*

[1] “Headlines should be as short as possible”. Newspaper headlines are, quite obviously, very short clauses. The actual length of each particular headline, however, is a matter of considerable debate and negotiation between senior editors and copy editors: Copy-editors, especially the beginners, suggest longer headlines, attempting to ‘capture’ as much of the story as possible. The senior editors shorten the headlines to a considerable extent- leaving out whole chunks of information. One of the expertises mastered by experienced editors is the ability to decide which parts of the story should be left *out* of the headline. The following exchange is a very typical example. The

<sup>9</sup> I thank an anonymous referee for his/her discussion of this point. The referee also wondered whether any of the ten properties may be reducible to another. Thus, for example, the referee felt that properties [6] and [7] are mirror-images of each other, and should thus be put together. I assume that this can indeed be done. For me, however, the more important point was that the editors I talked to felt these were two separate, although obviously related, principles. As I am interested here in the description of intuitions, rather than in the construction of a theoretical framework, I will discuss the two principles separately.

story is about a youth gang which was caught red-handed mutilating gravestones in a military cemetery in Haifa. The copy-editor suggested the following headline:

- (8) Haifa: A youth gang  
was caught mutilating  
gravestones in the  
city's military cemetery

The head of the news-desk ordered the copy-editor to shorten the headline, in the following way:

- (9) Haifa: A youth gang was  
caught mutilating gravestones

Note that the decision to shorten the headline is not without its price: We have lost a piece of information, i.e., that the gravestones were mutilated in a military cemetery, which means we have lost some contextual effects. We have, however, gained in reading effort. The shorter headline is simply easier to read. This is a very clear example of relevance optimization by effort reduction. The editor in charge decided that the loss in contextual effects is *smaller* than the gain in reading effort.

[2] “Headlines should be clear, easy to understand, and unambiguous”. Formulating a headline to a complex story is not an easy task. Copy-editors sometimes suggest headlines which come out unclear, difficult to understand, or unintentionally ambiguous. Such headlines are rejected, and the copy-editor is asked to formulate a clearer, simpler, unambiguous headline. In the following example, the article tells the story of a police drama in the city of Ramat-Gan, where a single arsonist threatened the city for weeks, burning down vehicles every night. On that specific night, the police caught a suspect, but had to release him after the ‘real’ arsonist took out to the streets again, burning down more vehicles to prove that he was not caught. The copy-editor suggested the following headline:

- (10) The ‘real arsonist’ from Ramat-Gan  
proves: You Haven’t caught me

The headline was rejected because it was considered unclear and unnecessarily ambiguous. It raises more questions and vaguenesses than it actually answers: Who is the ‘real arsonist’? Is there an ‘unreal arsonist’? How has the ‘real arsonist’ proven that he wasn’t caught? By whom? The copy-editor was ordered to formulate a clearer headline. This was his second attempt:

- (11) The arsonist ‘was caught’- and the  
vehicles in Ramat-Gan went on burning

This version is much clearer: It makes clear that a claim was made that the arsonist was caught, which turned out to be false, and it makes clear that, on that day,

after the capture incident ended, some vehicles in Ramat-Gan were still burning. There is, however, a potential ambiguity here: The headline can be read as making the claim that the arsonist set the vehicles on fire *before* the police made the false claim, and that they went on burning after the incident. The headline was rejected again, and the copy-editor suggested the third version, in (12), which was finally accepted and published as it is. Note that the only difference between (11), the rejected headline, and (12), the approved one, is in the tense of the verb in the second clause.

- (12) The arsonist ‘was caught’- and the  
vehicles in Ramat-Gan go on burning

This headline makes it clear that the arsonist is still on the loose, and is still in the habit of setting vehicles on fire. Finally, the story is captured in a clear, simple and unambiguous fashion. This reduces processing effort to the necessary minimum—and optimizes the relevance of the story.

[3] “Headlines should be interesting”. This quality plays a central role in the negotiations between copy-editors and senior editors. Many suggested headlines are rejected on the grounds that they are ‘not interesting’. What is usually meant by this rather obscure phrase is that the editor imagines that the *readers* of the paper will not find the headline interesting enough. In terms of our relevance-based theory, this means that the editor estimates that the amount of contextual effects carried by the headline will not justify the amount of reading effort. The copy-editor is then asked to read the article again, and look for a ‘more interesting’ piece of information to foreground to the headline. In the following example, the story includes an interview with Uri Lubrani, IDF’s Chief of Military Operations in Southern Lebanon. General Antoin Lahed, who is mentioned in the rejected headline, is the Commander in Chief of the South Lebanon Army (SLA), a Christian Militia which has traditionally been IDF’s ally in Lebanon. The context of the story is a wave of rumors, according to which the IDF plans to withdraw from Southern Lebanon, thus leaving General Lahed and his people on their own against their Islamic rivals:

This is the headline which the copy-editor suggested:

- (13) Lubrani: ‘There was no secret  
meeting with General Lahed’

Whether or not there was a secret meeting between IDF officials and General Lahed on the previous day is hardly an interesting question. After all, IDF officials and General Lahed meet on a regular basis, and their meetings are usually kept secret. In our terms, the headline does not carry a substantial amount of contextual effects. The headline was rejected, and the copy-editor came up with the following alternative:

- (14) Lubrani: ‘There is no plan to  
evacuate SLA seniors to Europe’

Whether or not there is a secret plan to evacuate SLA seniors to Europe is very obviously much more interesting. If there was such a plan, this would be a pretty



remarkable sign that the IDF is on its way out of Lebanon. Lubrani's flat denial can be interpreted in more than one way: We can take him for his word, or assume that he chose to deny the existence of the plan for tactical reasons— at any rate the denial has interesting implications. It is definitely more interesting than just another meeting— in our terms, it carries more contextual effects for the same amount of processing effort. Note that this is a very good example of the significant role of headline writing in the workings of a newspaper. The two headlines, the rejected one and the suggested one, make it quite obvious that the Lubrani interview did *not* contain any remarkable scoops, and that the editor had to dig in to find something which was worth promoting to the headline. As the senior editor's decision makes clear, even negative statements, flat denials of the type that Lubrani suggested as answers to the reporter's questions, have different amounts of relevance, and the one which was more relevant than the other was promoted.

[4] “Headlines should contain new information”. A major topic for negotiations between copy-editors and senior-editors has to do with the question of whether the readers already know what the copy-editor decided to promote to the headline. Obviously, editors do not really know what their readers know, but their *estimates* of their readers' state of knowledge play a central role in the decision-making process. This makes perfect sense within our relevance-based framework: A headline which does not contain novel assumptions cannot bring about contextual effects, and is thus irrelevant. In *Ma'ariv*, as in any other daily newspaper, estimates of the readers' state of knowledge are based primarily on what has already been communicated by the other mass-media, especially the evening news on TV. If the content of the proposed headline for the next morning has already appeared in the news the night before, most chances are it will be rejected. The following headline, for example, was rejected on these grounds:

- (15) The Austrian Chancellor  
Arrived for a visit; will  
meet Netanyahu today

The copy-editor had a hard time finding an alternative headline. This is what he came up with:

- (16) Officials in Jerusalem hope  
for the Austrian Chancellor's  
visit to run smoothly

This headline was accepted, for two reasons: First, it carries the implication that officials in Jerusalem are worried that the visit might *not* run smoothly—an angle on the visit which *was* new. Second, it connects the story to prior events and expectations: The visit of the British foreign minister had just ended the day before, and *that* visit was full of political hurdles and diplomatic embarrassments. As we shall see below, connecting a story to its wider context is another important property of good headlines.

[5] “Headlines should *not* presuppose information unknown to the readers”. This principle, in a sense, is the mirror-image of the previous one: The information in the headline should definitely be new—but it cannot be ‘overly new’. Headlines should only presuppose information which is already part of the *mutual knowledge* established between the newspaper and its readers. In terms of our relevance-based framework, every presupposition in the headline should already be available within the readers’ context of interpretation. Otherwise, the computation of the headline will result in zero contextual effects.

Consider, then, the following headline:

- (17) Advanced negotiations on the establishment  
of the second Israeli-owned casino in Jericho

This headline presupposes the existence, or at least the potential existence, of the *first* Israeli-owned casino in Jericho. The first news concerning the plans to build this casino, the first one, were published only a few days before the above headline was suggested. According to the editor in chief, the readers had not yet registered the future existence of the first casino in their long-term memory—it was premature to treat it as a presupposition. The copy-editor was asked to change the headline, and came up with the following alternative:

- (18) The first casino in Jericho  
will be operational in February

In this headline, the establishment of the first casino in Jericho is not presupposed, but reported as part of the news. This is much better. But the editor in chief asked the copy-editor to rephrase the headline again, this time for a different reason: The proposed headline forces the reader to *calculate* the amount of time it will take till the casino will be operational. This adds to the processing effort. The alternative, which was eventually published, reduces this effort, thus optimizing the relevance of the story:

- (19) The first casino in Jericho  
will be operational in a year

[6] “Headlines should include names and concepts with high ‘news value’ for the readers”.<sup>10</sup> Experienced editors develop a sense of the ‘news value’ of names and

<sup>10</sup> This property, and the next one, reminded one anonymous referee of Ariel’s (1988, 1991) accessibility theory (see also Kronrod and Engel 2000). The resemblance, however, is rather superficial. Ariel is not interested in the specific *contents* of the referring expressions, but in their general cognitive and structural properties. Thus, for example, referring expressions which function as high accessibility markers (e.g., personal pronouns, first names) are used by speakers when they assume that the referents are highly accessible for their addressees; lower accessibility markers (e.g., long definite descriptions, full names etc.) are used when the speaker assumes low mental accessibility for their addressees. The point I am making here, however, is not that newspaper editors prefer certain *types* of referring expression (e.g., full names) to others (e.g., last names), but that they prefer certain *referents* (e.g., famous figures) to others.

concepts: They very easily identify names and concepts which should appear in headlines- and those which do not. In terms of our relevance-based theory, experienced editors know, or at least believe they know, which names and concepts will carry a large number of contextual effects for their readers. The following example demonstrates this very clearly. Some background: Two days before the following story was to be published, a story in one of the national newspapers revealed that the popular musicians contracted to perform in Israel's 50<sup>th</sup> Jubilee were paid high sums of money- at the tax-payers' expense. In our story, some other popular musicians reacted to the revelation and angrily declared that they were willing to perform in the Jubilee for free. This was the headline suggested by the copy-editor:

- (20) A group of artists suggests  
an alternative for the Jubilee:  
“we are willing to perform for free”

This headline, a classic summarizing headline, was rejected in favor of the following, which replaces the expression ‘a group of artists’ with the names of two celebrity musicians, Shimi Tavori and Margalit Tsan’ani:

- (21) Shimi Tavori and Margalit  
Tsan’ani: “we are willing to  
perform in the jubilee for free”

Note, first, that the replacement headline in (21) actually loses some of the information we had in (20): The group of artists included many more musicians than just the two mentioned in (21). The point, however, is that the two are the most famous members of the group, and names of well-known popular musicians always carry a lot of contextual effects: This is so, because their names direct the readers to construct a much wider *context* for interpretation, which includes whatever we know about *them*, their personalities, their views, their social background, their wealth, their life styles, and so on. (20) is a headline suited for a regular news story about the 50<sup>th</sup> Jubilee; (21), on the other hand, is a headline for what is basically a *gossip* story about Shimi Tavori and Margalit Tsan’ani.

[7] “Headlines should *not* contain names and concepts with low ‘news value’ for the readers”. This is the mirror-image of the last property: Some names and concepts do not have “news value” for the readers, and experienced editors avoid promoting them to the headline. In our terms, these names and concepts do not help the reader construct the optimal context for the interpretation of the headline. In the following example, the copy-editor attempted to promote such a name to the headline, and was intercepted by the editor in charge. Some background: The story has to do with an accident in which two military helicopters collided in the air on their way to Israel's northern border, an accident in which more than 70 soldiers were killed. New immigrants to Israel receive an immigration grant from the Jewish agency, and it turned out that the agency asked the parents of one of the soldiers killed in the accident, a newly-arrived immigrant, to pay back his grant. The copy editor suggested the following headline:

- (22) The Michaelov family was ordered  
to return the immigration grant given  
to their son- who was killed in the Galilee

The editor in charge rejected the headline, and ordered the copy-editor to produce another one, which would not contain the name. This was the alternative headline, which was finally published:

- (23) The Jewish agency refused to let  
a family, whose son was killed in the  
helicopter accident, keep his immigration grant

Note that the move from the definite ‘the Michaelov family’ to the indefinite ‘a family’ made it difficult for the copy-editor to keep it in subject position, and dictated an overall grammatical change from passive to active voice.

[8] “Headlines should ‘connect’ the story to previously known facts and events”. Just like the last two principles, this one has to do with the construction of the appropriate context for interpretation. A story interpreted on its own, as an isolated event, will carry a certain amount of contextual effects. The same story can carry more contextual effects to the extent that the readers interprets it within a wider context, which includes previously known facts and events. Consider the following example, which is a report on a violent taxi robbery in the city of Haifa:

- (24) The driver was beaten  
and thrown out- and  
the stolen taxi was later  
found stuck in the mud

The editor in charge, who rejected this headline, asked the copy-editor to connect the incident in Haifa to the rising trend of taxi robbery throughout Israel. The following headline, which connects the specific incident in Haifa to the new criminal pattern, directs the reader to construct a context for interpretation which includes the former robberies, and promises to carry more contextual effects in this wider context:

- (25) Another taxi robbery:  
A driver from Haifa was  
attacked and thrown out  
of the vehicle

[9] “Headlines should ‘connect the story’ to prior expectations and assumptions”. Just as headlines have to connect the story to previously known facts and events, they have to connect the story to non-factual mental representations, i.e., prior expectations and assumptions which the readers may have with respect to the relevant topic. To the extent that the headline manages to do that, it helps the readers construct a context for interpretation in which more contextual effects will be

deduced. The following example needs no background: It has to do with one of the stages of the Lewinsky affair. This is the headline suggested by the copy-editor:

- (26) Clinton plans to admit that  
he kissed Monica Lewinsky

The editor in charge rejected the headline, and asked the copy-editor to suggest a headline which connects the story to the common assumptions and expectations which readers at that time had with respect to the story: The really important question was whether Clinton was about to admit to intercourse, not to kisses, with Lewinsky, and the assumption was that such a move on Clinton's behalf would constitute a very dramatic turning point in the whole affair, potentially leading to Clinton's impeachment. Connecting the story to this set of assumptions and expectations helps the reader to construct a more appropriate context for interpretation, and thus deduce a larger number of contextual effects. This is the headline which was finally published:

- (27) Clinton plans to admit to  
kisses- not to intercourse

Note that this headline does not contain any additional positive information: It merely explicates what the reader may have figured out from (26) by implicature. What it does do is position the story within its proper context, thus optimizing its relevance.

[10] "Headlines should 'frame' the story in an appropriate fashion". Many of the negotiations between senior editors and copy-editors have to do with the proper *framing* of the story. The characteristic question is: What kind of story is this? Is this, for example, a politics-oriented story, a human interest story, an entertainment story? As everybody who has ever worked with journalistic materials knows, the answer to these questions does not lie in the objective world, but in the construction of the story by its writer and editor. The following example has to do with the helicopter accident mentioned above. The copyeditor initially framed the story as a military-oriented story:

- (28) The defense ministry decided:  
The word 'disaster' will not be  
written on the gravestones of  
the victims of the helicopter accident

The editor in charge rejected this headline and asked the copy-editor to frame the story as a human interest one. According to the senior editor, framing the story as a human interest story would make it 'more interesting'—in our terms, it would help the reader construct a context for interpretation where the story would carry more contextual effects for the same processing effort. This is what the copy-editor came up with:

- (29) The parents' request was denied:  
The word 'disaster' will not be  
written on the gravestones of  
the victims of the helicopter accident

To sum up: The above list of properties constitutes an accurate rendition of the set of implicit professional imperatives shared by senior editors in *Ma'ariv*. As we have seen, each of these properties is reducible to a relevance-oriented strategy. Headlines can optimize relevance by requiring the *minimal amount of processing effort*- by being short, clear, unambiguous and easy to read. Headlines can optimize relevance by carrying the *maximal amount of contextual effects*- by being interesting and new. Headlines can optimize relevance by making sure the readers construct *the right context for interpretation*, and by making sure that their content is compatible with that context—by avoiding unknown presuppositions, by containing names and concepts with a high ‘news value’, by avoiding names and concepts with low ‘news value’, by connecting the story to previously known facts and prior expectations, and by framing the story in the proper fashion. As we have seen, headlines do not meet these criteria all at once. The art of headline production consists of formulating the headline which meets the maximal number of the above conditions, thus providing the reader with the optimal ratio between contextual effect and processing effort.

## 6. The strategy of tabloid headlines

From the point of view developed in this paper, tabloid headlines are not that different from the regular headlines which we find in more ‘respectable’ newspapers. Tabloid headlines simply take one relevance-optimization strategy, which we have already looked at, to its logical extreme. As we have seen before, headlines may produce more contextual effects by directing the reader to the appropriate context of interpretation. This is sometimes done at the expense of new information. In (25), for example, the fact that the stolen taxi was later on found stuck in the mud was demoted from the headline, and the expression ‘another taxi robbery’ was added, in order to instruct the reader to retrieve information from long-term memory concerning previous taxi robberies- thus constructing the optimal context for interpretation. Tabloid headlines can be thought of as adopting this strategy all across the board: Keep your processing effort *and your new information* to the minimum, and optimize relevance by maximizing the context of interpretation. This is actually a theoretical re-interpretation of Lindemann’s formulations: Tabloid headlines are not very informative, but they very efficiently *trigger frames and belief systems in the reader’s mind; they evoke images and scenarios in the reader*. Let us look at some of Lindemann’s examples:

- (30) a. NO-LA-LA! The Frogs Get Bored with Bed  
 b. ‘Dirt’ at Posh Noshery  
 c. Boy’s Whisky Hell  
 d. Space Ape Makes A Monkey Out of Moscow

The first thing to note is that although these headlines are relatively uninformative, none of them is *completely* devoid of new information: (30a) reveals that the French are becoming bored with sex; (30b) says that something dirty was discovered in an exclusive restaurant; (30c) informs that some boy went through an

ordeal involving whisky; and (30d) asserts that a space ape drove Moscow crazy. These new pieces of information may not have a very clear designation, and we may be left with a lot of unresolved questions—*which restaurant? What dirt? Which ordeal? What is a space ape anyway?*—but we nevertheless get a minimal amount of new information out of each of these headlines.

The second and crucial point is that each of these headlines very efficiently instructs the reader to construct an extremely rich context for interpretation—a context full of clichés and prejudices, and feelings of fear, passion and hatred—in which even the informationally-dull headlines carry more contextual effects than a lot of informationally-rich headlines in the more respectable newspapers. Consider the headline in (30b). As its story reveals, a top restaurant in London faced allegations of breaching hygiene regulations. In a regular newspaper, the same story would probably be published under a headline like ‘Top Restaurant Faces Hygiene Allegations’. This headline is slightly more informative than (30b), but it does not get anywhere close to the amount of contextual effects produced by (30b). The tabloid headline very efficiently raises in the readers’ minds a complex set of notions—vivid images of dirt and filth, feelings of envious contempt towards the rich people who can afford to eat in fancy restaurants, and so on and so forth—which then constitute the context for the interpretation of the headline. In this context, the slight information in the headline carries a great deal of contextual effects: The story is no longer that of the specific restaurant in London, but a generalized story “revealing” the “real” dirt behind the “facade” of expensive cuisine and the life-style of the rich and famous. The same is true for the other headlines as well: (30a), as we have already seen, raises a context consisting of a mixture of sexual frustration and hatred and contempt towards the French. (30c) raises a context for interpretation consisting of all the generalized fears involving children, alcohol and violence, and thus carries a large set of contextual effects: The headline translates a story about *one* specific child into a story about children in general and their fragile existence—and the context constructed by the readers’ obviously includes their *own* children. (30d) makes use of the mixture of fear and contempt towards the Soviet Union, and ridicules it with a clever combination of reality and metaphor. The contextual effects, again, go well beyond the fact that the specific monkey sent to space by the Soviets was interfering with the space mission; the tabloid headline is now about the “incompetence” of the fearful Soviet Union, a whole super-power ridiculed by a single monkey.

## 7. Do headlines really attract readers to their stories?

As we have seen in [section 2](#), it has been claimed in the literature that newspaper headlines have an additional function, which goes beyond the strict semantic-referential function, namely to “attract the attention of the reader and provoke the reader to read the whole story” (Nir, 1993). In this section, I will discuss this function, and claim that it is better characterized as a relevance-based *selection* mechanism, which rather than indiscriminately ‘attract’ all readers to all the stories in the

newspaper, serves to direct *individual* readers to the *subset* of stories which would carry the optimal relevance for *them*.

The empirical starting point for this discussion is the rather obvious fact that readers do not always read news items beyond the headline. On the contrary, most readers spend most of their reading time scanning the headlines than reading the stories. As Nir (1993) notes, “for the modern newspaper reader, reading the headline of a news item replaces the reading of the whole story” (p. 24). If we stick to the above definition of headlines as ‘attracting devices’, we are forced to conclude that most headlines fail in fulfilling their function— a rather unintuitive conclusion to say the least: If that was indeed the case, the headline would not survive the dynamic evolution in the design of the modern newspaper. A much more reasonable assumption is that the characteristic reading pattern manifested by scanning readers is *exactly* what the headline is supposed to achieve. Let us see why this is the case:

The major claim made in this paper has been that headlines are relevance-optimizers. This characterization of the headline means that an ordinary reader who has finished reading a headline has *already* received the optimal amount of relevance for its story. This means that reading beyond the headline, through the whole text, would actually amount to a process of gradual *reduction* of the relevance of the story for the reader. Think about the Kennedy story we have looked at before. The first sentence of the story—*The bodies of John Kennedy Jr., his wife Caroline and his sister-in-law Lorraine were discovered yesterday in the ocean, at a depth of 30 meters, 10 kilometers away from Martha’s Vineyard island, where they were headed on Saturday*— adds considerably to the processing effort, and adds some new information. The crucial point, however, is that the additional amount of new information does not necessarily add a sufficient amount of contextual effects to make it worth the ordinary reader’s while to go through the interpretation process. The fact that the bodies were found at a depth of 30 meters, for example, would probably have no real contextual effect for the ordinary reader. The second and third sentences—*Senator Edward Kennedy, John’s uncle, arrived at the site where the bodies were found, in order to identify them. Kennedy Jr. will be buried in NY in the coming days*— add to the processing effort, but have a meager contextual effect. All this means is that the reader who decides to read the headline—*John Kennedy Jr.’s body found*— and move on to the next headline, rather than delve into the story, makes a perfectly rational decision, and actually gets the best (informational) value for (cognitive) money possible. The selection of the best headline for a story is thus not supposed to make the ordinary reader go on reading the story, but to insure that the reader has indeed received the best ‘deal’ in *reading the headline itself*.

Does that mean that readers who nevertheless go on reading the whole story make an irrational decision? Obviously not. According to our relevance-based conception, we should expect two types of readers to go on reading the story beyond the headline— and still be rational in doing it. First, a fair number of readers may have a good reason to expect to receive *more* contextual effects from the story than the ordinary reader. Among these, we should find, for example, Kennedy aficionados, celebrity enthusiasts or aviation experts. Note that in regular speech, we would simply say that these specific readers have a special *interest* in the story. In relevance-theory



jargon, we say that the cognitive context which they construct for the interpretation of the story allows for the deduction of more contextual effects for the same amount of processing effort. If an individual reader already knows a lot about Caroline Kennedy, for example, it would make sense for him or her to go on reading through the story, because her death would carry a significant amount of contextual effects for that reader. In this sense, the headline serves as a *selection-device* for the readers, directing each individual reader to those specific stories which may justify the investment of additional cognitive effort in providing additional contextual dividends.

The second type of readers who would probably go on reading the story to the end are those who would be willing to put in the extra effort *even* if the contextual effects would not justify that. Among those, we should find the avid readers- who enjoy spending time reading a newspaper regardless of the specific contents of the stories. In terms of relevance-theory, those readers have a different cognitive style than the ordinary reader- the threshold they set for the ratio between contextual effects and processing effort is lower than that of the ordinary reader. Note that cognitive styles form a continuum: Some people, for example, may find the computation of the same contextual effect more intellectually demanding than others; for those, the relevance of any new assumption will be reduced in proportion to their processing effort. Specifically, other things being equal, proficient readers will find more news stories to be relevant than less-proficient ones. Moreover, some people may be generally more curious, interested or cognitively-energetic than others; they will be willing to spend more energy for the computation of the same contextual effect. Other things being equal, these people will be willing to spend more energy on the same news story than the others, regardless of the amount of contextual effects carried by the story. This continuum of cognitive styles plays a significant role in editorial policies: For example, ‘quality’ newspapers regularly assume that their readers are more proficient, cognitively-energetic and curious than ‘popular’ newspapers take their readers to be. This explains, for example, the fact that headlines in ‘quality’ newspapers are, statistically speaking, longer, more complex and more difficult to read than headlines in ‘popular’ newspapers. Note, moreover, that in those rare cases when we set the threshold for the ratio between contextual effects to processing effort to *zero* - headlines lose their functional role. This happens, for example, when we read every word in an old newspaper in order to kill time waiting for a flight. In this extreme context the headlines become *redundant*: If we make a decision to read the whole newspaper, word by word, we no longer need the headlines to direct us to those stories which are best suited for our interests.

## 8. Conclusion

Newspaper readers are flooded on a daily basis with an amount of new information which they have neither the time nor the energy to process. Newspaper headlines help them get the maximum out of this informational flood—for the minimal cognitive investment. First, headlines provide the readers with an optimally relevant presentation of their stories. A good headline is one which helps the reader deduce the maximal amount of contextual effects for the minimal amount of processing

effort. Then, they guide individual readers to those specific stories which would be worth their while to read in the full version. Thus, the reading patterns manifested by newspaper readers are exactly what we should expect: Readers regularly scan the headlines, and only occasionally stop to read the actual story. As relevance-optimizers and relevance-based selection-devices, headlines function as *negotiators* between stories and readers. As we have seen, producing the appropriate headline for a story is a complex task exactly *because* the headline is neither a semantic summary of the story nor a pragmatic attracting-device for the reader, but a communicative device whose function is to produce the optimal level of affinity between the content of the story and the reader's context of interpretation, in order to render the story optimally relevant for the reader. As the analysis of headline production in *Ma'ariv* clearly shows, this delicate process involves the constant juggling of a large number of different, and sometimes contradictory, communicative imperatives.

As we have seen, our relevance-based conception of newspaper headlines has allowed us to go beyond the descriptive distinctions between different types of headlines, and *explain* their functions as tactical variations on the theme of relevance optimization. It thus provides for a novel way to conduct comparative analyses of headlines in different types of *newspapers* and in different types of *cultures* of communication. As I have already suggested, many of the major differences between newspapers may be attributed to the assumptions made by the editors concerning the set of relevance considerations shared by their readers—their cognitive styles, their reading proficiency, their interests, their views, their fears and passions, and so on and so forth. Moreover, it is an intriguing question, and as far as I can tell, a totally open one, whether different newspapers are actually successful in *predicting* the relevance-oriented profiles of their readers. I would venture to hypothesize that those newspapers which are successful at that are those which maintain a steady and well-defined readership.

Finally, on a larger scale, the analysis presented in this paper should be thought of as an exercise in the application of universalistic, psychologically-oriented theories to the explanation of culturally-specific, variable social phenomena. Modern newspaper headlines are a very late cultural development, yet their functional nature is best characterized using a set of constitutive, cognitive notions which are applicable in principle to a very wide set of acts of intentional communication. As opposed to some of the current thinking in the social sciences, the imposition of a universalistic theory on a specific cultural phenomenon does *not* automatically entail indifference to the fine-grained patterns of social variability involved. In our case, for example, the universalistic framework actually helps us capture the differences and similarities between headlines in regular and tabloid newspapers in an *explanatory*, rather than a merely descriptive, fashion. I take this to be a very encouraging sign that the importation of theoretical insights from the cognitive domain into the social sciences is a worthwhile endeavor.

### Acknowledgements

I thank two anonymous referees for their helpful and insightful comments on this paper.

## References

- Ariel, Mira, 1988. Referring and accessibility. *Journal of Linguistics* 24, 65–87.
- Ariel, Mira, 1991. The function of accessibility in a theory of grammar. *Journal of Pragmatics* 16, 443–463.
- Bell, Allan, 1984. Good copy, bad news: the syntax and semantics of news editing. In: Trudgill, P. (Ed.), *Applied Sociolinguistics*. Academic Press, London.
- Bell, Allan, 1991. *The Language of News Media*. Blackwell, Oxford.
- Bernstein, M., Garst, R.E., 1982. *Headlines are Deadlines*. Columbia University Press, New York.
- Chang, Tsan-Kuo, Pollick, Barry, Lee, Jae-Won, 1992. The world as subjective reality: US newspaper editor's view and its determinants. *International Journal of Public Opinion Research* 4 (2), 177–183.
- de-Knop, Sabine, 1985. Linguistic and extralinguistic aids for reconstruction and interpretation of metaphors in headlines. In: Paprotte, W., Dirven, R. (Eds.), *The Ubiquity of Metaphor: Metaphor in Language and Thought*. Benjamins, Amsterdam.
- Dierick, Jean, 1987. Equivalent messages: headlines in *Le Monde* and their translation into English. In: Maes-Jelinek, H., Michel, P., Michel-Michot, P. (Eds.), *Multiple Worlds, Multiple Words*. University of Liege, Liege.
- Dor, Daniel, 2001. *Newspapers Under the Influence*. Babel Publications (in Hebrew).
- Fasold, Ralph, 1987. Language policy and change: sexist language in the periodical news media. In: Georgetown University Round Table on Languages and Linguistics (GURT), 187–206.
- Galtung, Johan, Ruge, Mary Holmboe, 1965. The structure of foreign news. *Journal of Peace Research* 2 (1), 64–91.
- Henley, Nancy M., Miller, Michele, Beazley, Jo Anne, 1995. Syntax, semantics, and sexual violence: Agency and the passive voice. *Journal of Language and Social Psychology* 14 (1–2), 60–84.
- Iarovici, Edith, Amel, Rodica, 1989. The strategy of the headline. *Semiotica* 77-4, 441–459.
- Jenkins, Helen, 1990. Train sex man fined: headlines and cataphoric ellipsis. In: Halliday, M.A.K., Gibbons, John, Nicholas, Howard (Eds.), *Learning, Keeping and Using Language: Selected Papers from the 8th World Congress of Applied Linguistics*, Sydney. Benjamins, Amsterdam, pp. 16–21.
- Kronrod, Ann, Engel, Orit, 2001. Accessibility theory and referring expressions in newspaper headlines. *Journal of Pragmatics* 33, 683–699.
- Leon, J.A., 1997. The effects of headlines and summaries on news comprehension and recall. *Reading and Writing* 9, 2. 85–106.
- Lindemann, Bernard, 1989. What knowledge does it take to read a newspaper? *Journal of Literary Semantics* XVIII/1, 50–65.
- Lindemann, Bernard, 1990. Cheap thrills we live by: some notes on the poetics of tabloid headlines. *Journal of Literary Semantics* 19–1, 46–59.
- Mardh, L., 1980. *Headlines: On the Grammar of English Front Page Headlines*. CWK Gleerup, Lund.
- Nir, Raphael, 1993. A discourse analysis of news headlines. *Hebrew Linguistics* 37, 23–31 (in Hebrew).
- Perfetti, Charles A., Beverly, Sylvia, Bell, Laura, Rodgers, Kimberly, Faux, Robert, 1987. Comprehending newspaper headlines. *Journal of Memory and Language* 26 (6), 692–713.
- Pfau, Michael R., 1995. Covering urban unrest: the headline says it all. *Journal of Urban Affairs* 17 (2), 131–141.
- Sidiropoulou, Maria, 1995. Headlining in translation: English vs. Greek press. *Target* 7 (2), 285–304.
- Sperber, Dan, Wilson, Dierdre, 1986. *Relevance: Communication and Cognition*. Blackwell, Oxford.
- van Dijk, Teun A., 1988. *News as Discourse*. Lawrence Erlbaum Associates, Publishers.

**Dr. Daniel Dor** teaches at the Dept. of Communication and the Dept. of English, Tel Aviv University. His research interests include, among other topics, the role of the mass media in the construction of political hegemony, the linguistic consequences of globalization, and the cultural-biological evolution of language.