Entry points and reading paths on newspaper spreads: comparing a semiotic analysis with eye-tracking measurements

JANA HOLSANOVA AND KENNETH HOLMQVIST
Lund University, Sweden
HENRIK RAHM
University College of Kristianstad and Lund University, Sweden

ABSTRACT
The aim of this article is to compare general assumptions about newspaper reading with eye-tracking data from readers’ actual interaction with a newspaper. First, we extract assumptions about the way people read newspapers from socio-semiotic research. Second, we apply these assumptions by analysing a newspaper spread; this is done without any previous knowledge of actual reading behaviour. Finally, we use eye-tracking to empirically examine so-called entry points and reading paths. Eye movement data on reading newspaper spreads are analysed in three different ways: the time sequence in which different areas attract attention is calculated in order to determine reading priorities; the amount of time spent on different areas is calculated in order to determine which areas have been read most; the depth of attention is calculated in order to determine how carefully those areas have been read. General assumptions extracted from the socio-semiotic framework are compared to the results of the actual behaviour of subjects reading the newspaper spread. The results show that the empirical data confirm some of the extracted assumptions. The reading paths of the five subjects participating in the eye-tracking tests suggest that there are three main categories of readers: editorial readers, overview readers and focused readers.

1. INTRODUCTION
For many people, it is a matter of habit to read a newspaper in the morning and they usually concentrate on particular sections and pages. But it is not at all self-evident how we actually read newspapers. What do we choose from a newspaper spread? Which items and areas of the spread do we attend to and which of these do we actually read? Are reading paths influenced by the newspaper layout?
In this article, we examine two ways of analysing a newspaper spread in the Swedish morning paper *Svenska Dagbladet*. The purpose is to compare results from a semiotic analysis with empirical results from readers' actual interaction with newspapers analysed by eye-tracking measurements.

We present two different analyses of this newspaper spread. First, we approach the text *deductively* through a semiotic analysis. In the second step, we present the result of eye-tracking measurements on the spread, that is, *inductively*. Finally, we discuss the results of the two methods of analysis. We would like to stress that the semiotic analysis has been made without any previous knowledge of the results of the eye-tracking measurements.
For our analysis, we chose a newspaper spread from *Svenska Dagbladet*, 5 September 2002, pp. 16–17, see Figures 1(a) and (b).

2. A SEMIOTIC APPROACH

Semiotic research has provided models for analysing text, layout and illustrations in texts, in order to acquire an integrated and deeper understanding of the interplay between society, semiotics and text linguistics, for example the socio-semiotic works by Kress and Van Leeuwen (1990, 1996, 1998). A central interest is the understanding and interpretation of media
texts. Swedish researchers (even though they are not outspoken socio-semioticians) have analysed layout in terms of semiotic spaces (see later) and formulated assumptions on meaningful units and probable reading paths (Melin and Pettersson 1991; Melin, 1995, 2000).

Taking a semiotic or socio-semiotic approach is a heuristic way of interpreting verbal and visual units. It is important to bear in mind that a semiotic analysis cannot provide an empirical truth – empirical in the sense of proving how verbal and visual units are interpreted. The aim is rather to contribute to the understanding of the more usual human ways of interpreting signs. Before carrying out our analysis, we will examine semiotic space.

2.1 Semiotic space and its context levels

Our starting point is to regard a newspaper spread as a semiotic space (Kress and Van Leeuwen, 1996). The concept of semiotic space implies that meaning is created through an interplay of verbal units and visual means of expression. In Swedish research, one of the most important studies of semiotic space is Ledin’s (2000) analysis of the Swedish weekly press of the 19th century.

It is important to discuss not only how verbal and visual units create meaning, but also which actors are involved in the semiotic process. In briefly describing these actors later on in this article, we try to contextualize semiotic space by discussing different context levels. Thus, we are going beyond merely describing semiotic space.

In the context of the media, there are two main actors: the producers and consumers of the newspaper spread. At the most concrete level, the producers are the journalists writing the texts, the photographers taking the photographs and the designers of other artwork. On the next level the newspaper layout artists combine the verbal and visual units into articles and add commercials to make up a newspaper spread. Moving up another level in the producer hierarchy, we come to the in-house expectations of the actual newspaper offices (here Svenska Dagbladet) of how to design a spread in a certain section of the newspaper (in our case the domestic news section). Finally, there are common traditions, ideals and expectations within the journalistic discourse of how to create meaning.

The semiotic expectations of these different levels of journalistic producers are interconnected to different levels of the consumers of journalistic products, in this article the viewers of the spread. The most concrete level is of course the single viewer of the spread. However, the journalistic product is not intended for just one type of viewer; the producers also bear in mind the various categories of single viewers, according to gender, age, ethnicity, ideology, education, interests, etc. Thus the producers of a major newspaper – like the one investigated here – need to consider many different kinds of single viewers as well as the average reader.
These interconnections between different levels and categories of producers and consumers have consequences for the semiotic analysis. The fundamental assumption for such an analysis is that there are tacit agreements on how to handle the creation and interpretation of meaning on a newspaper spread. The task of the semiotician is to consider these agreements when analysing the actual spread.

How is semiotic space connected to other semiotic situations? It is possible to compare the indirect communication between producer and viewer with the conditions for participants in direct communication. In direct face-to-face communication, participants understand each other by mutually observing and interpreting linguistic, paralinguistic and extra-linguistic modes of communicating. For both parties, there are possibilities of reformulating and reinterpreting messages. As mentioned earlier, the special form of indirect communication taking place between journalistic producers and consumers has to rely on tacit agreements. Kress and Van Leeuwen (1996) see this as a disjunction between producer and consumer where the ‘articulation and understanding of social meanings in images derives from the visual articulation of social meanings in face-to-face interaction’ (p. 119). According to Halliday (1978), this way of interpreting semiotics is the interpersonal function. The framework outlined by Kress and Van Leeuwen can also be applied to the two other metafunctions formulated by Halliday: the ideational and the textual.

However, the semiotic analysis in this article focuses on the textual (or compositional) metafunction for two main reasons: (i) the connection to other analytical tools as paratexts (see later); and (ii) the possibility of comparing a semiotic analysis of the composition of the spread with eye-tracking analyses of individuals interpreting the semiotic components of the spread.

2.2 Paratexts as entry points

The body of the various texts quantitatively dominate the verbal entities of a newspaper spread. However, at least as important as the main text are the paratexts. Paratext is a common notion for the different accompanying texts that connect to the main texts, originally defined by the literary scholar Gérard Genette (1997[1987]) who also described paratexts as thresholds into the text. Paratexts encompass headlines, subheadlines, picture captions, billboards, tickers (a very brief synopsis directly below the headline), but also paratexts of a more global kind, such as recurring vignettes, thematic markers or markers indicating to which section the pages belong (e.g. local, international, economy).

The paratext can also be seen as a connecting link between the producer and the consumer. It is derived from two semiotic systems: the verbal–linguistic and the typographical as paratexts are both verbal signs and non-verbal signs depending on choices of typography and layout (Frandsen, 1992: 153). Of course, journalists, editors and designers use paratexts in both
these functions. And the viewers will concomitantly interpret the paratexts in both these senses, making use of paratexts as entry points between the two semiotic systems as well as between semiotic units. Paratext can serve as starting point, as invitation to reading (cf. section 3.1). Paratext is thus a central notion for creating, recreating and interpreting the newspaper spread semiotics.

The notion of paratext has been applied in different scientific works in Sweden. An early example is Persson (1994) who uses it for analyses of the Swedish evening paper *Expressen*. Melin (2000) discusses paratexts in his description of how everyday non-fiction texts have changed since the definitive break-through of computers in the 1990s, in order to show the interplay between graphic and verbal means of expression. Rahm (2002) uses paratext to distinguish and analyse various genres in the daily press from the 19th century to the 1990s.

2.3 The meaning of composition and seven assumptions on reading paths

Before presenting the first analysis, let us return to Kress and Van Leeuwen (1996: 183ff) for a moment by briefly presenting their theoretical model for analysing the semiotics of composition. In this model, there are three main interrelated components: information value and structure, salience and framing. The information structure of the semiotic space follows three separate dimensions: the vertical dimension, the horizontal dimension and the dimension of the centre–periphery. In the vertical dimension, general information is found at the top graduating down to more specific information at the bottom. The horizontal dimension has ‘given’ information on the left and ‘new’ information on the right. The third dimension for information structuring can be seen as a way of organizing information by placing most important information in the centre and less important information outside the central position. But this dimension of centre–periphery can also function as an intermediary link between the other two. See also Rahm (in press) where these dimensions are some of the tools used in analyses of direct mail.

The second component deals with the degree of salience of the various semiotic units. Factors involved in salience are foregrounding–backgrounding, relative size, contrasts in tonal value or colour and differences in sharpness of focus.

The third component is framing and encompasses framing lines, dividing lines and connecting lines between semiotic units. Thus these last two components deal with the layout of the spread, especially non-verbal units and paratexts.

In order to compare general assumptions about reading behaviour with the results of our empirical investigations, we first had to search for them in socio-semiotic research. Although we did not find a ready-made
explicit list of assumptions, we were able to extract seven general assumptions from Kress and Van Leeuwen (1996):

1. Readers prefer new information and expect this to be on the right in the semiotic space.
2. Readers prefer the most general information at the top and the most specific information at the bottom of the semiotic space.
3. Readers look for the most important information in the centre of the page and less important information on the periphery.
4. Readers look for graphically salient elements; however, it is important to bear in mind that ‘what is made salient is culturally determined’ (Kress and Van Leeuwen, 1996: 219).
5. Readers look for paratexts.
6. Readers follow elements connected to each other by framing devices such as lines and arrows.
7. Readers scan the semiotic space before taking a closer look at certain units.1

Possible combinations of these assumptions are, for example, placing new information on the right of the semiotic space, presenting new information in the form of an eye-catching illustration and connecting this illustration to other semiotic units by framing devices such as lines and arrows, either explicit or implicit in the illustration itself. Kress and Van Leeuwen (1996) describe ‘the most plausible reading path’ being taken when readers ‘begin by glancing at the photo, and then make a new start from left to right, from headline to photo, after which, optionally, they move to the body of the verbal text’ (p. 219).

These seven general assumptions about reading paths are applied in the semiotic analysis in section 2.4, where they serve as a point of departure for the comparison of general assumptions and actual reading behaviour (section 4).

2.4 A semiotic analysis of the news spread

On a newspaper spread such as the one under analysis from Svenska Dagbladet (see Figure 1) with several texts about different topics, photos and other semiotic units, we assume that the reader starts by an overall scanning of the spread. (The reader we have in mind is a notional average reader, ‘average’ in terms of age, education, interests, etc.) As an average reader cannot be linked to specific interests in specific topics, we assume that none of the texts is preferred because of its subject matter.

In the upper left corner of Figure 1, there is a headline ‘VEM SKA HA PRISET?’ (Who will win the award?). This headline is very eye-catching because its font size and style (large white capitals against a red background) combine to grab the attention of the reader after the initial scan of the
newspaper spread. The combination of the two linear dimensions for information value would suggest that this is the place for general and known information. The definite form of the noun priset (the award) indicates that the information is given, but this is of course pseudo given information. The assertedly known information could work as a decoy for continued reading of the ad, which asks for nominations for awards in different journalistic fields. In this case, the information will become more specific as the reader continues towards the foot of the ad. The most specific information is the listing of people who are members of the award committee and the background information that the award existed since 1966.

But is it plausible to assume that reading always takes the same course? This can hardly be the case if the spread is looked upon as a semiotic unit, competing with many other spreads in the newspaper. The assumed overall scanning and the remaining attention to the spread is likely to be around a minute or, at best, maybe five minutes. A reasonable point of departure is that the graphically salient will attract attention in the first place. This would suggest a diagonal scanning in the direction of the bottom half of the right page, where we find the headline 'Studentilska mot bostadsbrist' (Student anger against housing shortage) and a photograph of a protest rally with participants carrying posters. In the ticker to the text, the reader is informed that the text addresses a typical news event for a daily – an event from the preceding day – as the text explicitly states that the rally took place yesterday. The position of the text on the lower part of the right page can be interpreted as new and specific information according to Kress and Van Leeuwen. The question is whether this affects reading as this topic is no newer or more specific than any of the other topics dealt with on the spread. A plausible interpretation is to study separate elements of the information structure in Studentilska mot bostadsbrist. Then it makes sense that the paratext in the form of the headline is most general, at the same time functioning as a decoy into the body of the news text as it does not say anything (indeed cannot say anything due to the limited number of letters in a headline) about which students are angry about which housing shortage in which town.

On the outer margin of the right page there is a column consisting of news items, i.e. a collection of short news texts. The only reasonable interpretation of this collection of texts is that they are read horizontally, since they are separate texts and have been placed in this column because they are short summaries of recent events. The paratexts are important for this column. At the top we find the global paratexts 'VAL 2002' (The elections 2002) and 'NYHETER' (News). The first one indicates the common theme for the column, the Swedish parliamentary election in September 2002. The paratext Nyheter (News), which also occurs on the preceding and the following spreads, is not only indicative of the column of news items but also signals that the entire spread belongs to the news section. In this column, the fourth text is graphically salient, which could mean that it easily catches the
reader’s attention when scanning the spread. There are several reasons for this: the text has a larger headline than the others; it is preceded by the paratext ‘Fråga & Svar’ (Q&A, which signals a regularly occurring and thus well-known section of the newspaper); and it is illustrated by a photograph of the interviewed person.

In the middle of the spread, there are two longer texts. Both should be read vertically with the most general information in the headline and the ticker. None of them are graphically salient, and it is therefore not safe to assume that they will be taken in during the first scanning when the reader establishes an overview of the semiotic space. If and when the reader pays more attention to the spread, the paratexts probably play an important role. Then the headlines of these two articles are important, as well as the global paratexts of the column of news items and headlines for single news items.

3. EYE-TRACKING AND NEWSPAPER READING

Data on visual behaviour can be used as measures of attentional processes. A relatively small number of explorative and experimental eye-tracking studies of newspaper reading have been conducted since the late 1980s. This section starts with a short overview of previous research in the field of studies of newspaper reading based on eye-tracking. This is followed by analyses of the results from empirical tests.

3.1 Previous research: entry points and reading paths

In the early eye-tracking studies (Küpper, 1989; Garcia and Stark, 1991), eye movements were recorded as videos. Video recordings are illustrative since we can follow the reader’s eye movements online. However, if we want to draw general conclusions, it is necessary to watch the video in slow motion and encode specific parameters frame by frame. In recent studies at the Lund Eye-Tracking Laboratory, eye-tracking was combined with head-tracking using a virtual reality model of the table and the newspaper. Subjects wore a helmet equipped with a head-tracking and eye-tracking device. No manual coding was necessary: the computer did the decoding based on absolute data co-ordinates representing the point of gaze on a specific plane. Figure 2 shows the experimental setting and equipment.

Garcia and Stark (1991) tested 90 readers of three newspapers in the USA. A major (and influential) finding in their study was that readers do not really read but rather scan newspapers. At certain entry points they stop scanning and start reading the story connected to the entry point. Garcia and Stark’s discussion ends by defining the task for newspaper designers: ‘to give readers material that is worthy of their scan, that makes them stop scanning and start reading’ (p. 67).

Hansen (1994) studied 12 readers of the Copenhagen newspaper Det Fri Aktuelt. His unpublished study investigates the order in which objects on spreads were scanned. He found that pictures are looked at first, then icons...
and graphics, followed by headlines of different sizes and text, with items in the usual newspaper format being observed last. Hansen also investigates readers’ priorities concerning article length, position and genre.

Apart from reading paths and entry points, there has been specific research interest in how ads are read. Widman and Polansky (1990) tested 129 readers of the Stockholm newspaper *Dagens Nyheter*. In a follow-up study, Lundqvist and Holmqvist (2001) tested 14 readers of *Dagens Nyheter*, focusing on the effect of size on attention, attitude and memory of ads.

More recent studies have focused on reading in new media, such as internet articles. Holmqvist et al. (2003) studied the proportion of reading and scanning in Swedish newspapers and internet articles. With the help of a reading filter, two distinct reader behaviours were identified. Readers of newspapers were found to read 55 percent of the time and scan 45 percent of the time whereas internet article readers read only 44 percent of the time and scanned 56 percent of the time. Entry points and reading paths, as well as orientation and navigation in Swedish internet articles (as compared to newspapers), were studied in Holsanova and Holmqvist (2004).

From a methodological point of view, virtually all eye-tracking studies of newspaper reading have been case studies. With the partial exception of García and Stark (1991), no one appears to have carried out controlled experiments. Hypotheses on which newspaper design factors attract readers’ attention have focused on the spread layout, size effects,
colour, drop quotes, fact boxes, pictures, axiality of objects, etc. In a large study of 17 Nordic newspapers, Holmqvist and Wartenberg (2005) presented results on how well designers predicted the probable effect of all these design factors. The results from reading and scanning paths are generalized at a group level. In a study of information graphics, Holsanova et al. (forthcoming) combine a naturalistic study design with controlled conditions.

In the following analysis, we concentrate on how newspaper readers are attracted by various areas of interest on the spread. In other words, we study which entry points they choose and trace their reading paths.

3.2 Analysis of results from eye-tracking measurements

The spread from Svenska Dagbladet under analysis in this article was used in a demonstration at the Society for Newspaper Design/Scandinavia Conference, in Malmö, Sweden, in September 2002. During the conference, one of the authors, Kenneth Holmqvist, installed a drop-in laboratory to investigate newspaper reading using eye-tracking. Kenneth Holmqvist and Constanze Wartenberg, who jointly operated the equipment, invited anyone interested to record their reading of a spread. Afterwards those taking part were offered the chance to see their eye movements and learn a little about research in this area. All readers were involved in some way or another in the newspaper world.

Each subject could choose 1 of 12 spreads; 5 people out of 17 chose to read the spread in Svenska Dagbladet analysed below. The spread was placed behind a calibration screen (a white board with 9 crosses), the subject sat down 80 cm from the spread, and the eye-tracker was calibrated. The subjects were asked to read as they did normally as soon as the calibration screen was removed. The subjects read the spread for as long as they wanted, while their eye movements were being measured.

What were the results of the eye-tracking measurements? A first impression of the actual paths across the newspaper spread can be seen in the scan paths in Figure 3, which are shown from the first minute of reading and scanning the spread. The circles indicate fixations, i.e. pauses when the eye rests for intake of information (on average for 170 ms). Larger circles indicate longer fixations. There is a 1-second filled reference circle in the bottom right corner. Lines indicate saccades, which are quick jumps (20–50 milliseconds), when the reader changed the position of the eye to reach the next landing position. During saccades, we are virtually blind since visual information intake is blocked.

The eye movements revealed the following pattern: newspaper readers do not read in the ordinary sense. Rather they scan the spread looking for entry points. When they have found an interesting entry point, they stop scanning and start reading for a while. After a while, scanning is resumed until the next interesting entry point is found. In other words, there is usually no phase of overview scanning preceding the actual reading; reading and
Figure 3: The first minute of newspaper reading of one of the subjects presented as scan paths.
scanning continuously overlap. The most common entry points are pictures and headlines, but drop quotes and fact boxes are also frequent entry points. The effect of colour is disputed, but appears not to have any strong influence.

In order to describe in detail what has been read first, most and most thoroughly, we performed three types of analyses based on the questions in the headings of the following subsections.

3.2.1 In what order were the different parts of the spread attended to?
The first analysis focuses on reading paths and reading priorities. The spread has been segmented into 10 areas (see Figure 4): ‘VEM SKA HA PRISET?’ [HEADLINE], ‘Hela journalistpriset’ [AD TEXT], ‘39-åringen, Två notiser’, ‘PIGGVAR’ [AD], ‘Koldioxidskatten’, ‘Studentilska’ [PHOTO], ‘Studentilska’ [TEXT], ‘Valspalten’, ‘Frihetsfeminism’. The areas concern the entire texts, headlines or pictures with picture caption. In one case, the area covers two short articles in a graphically clearly delimited space.

For each subject, we thus mapped the time sequence order in which these defined areas caught the attention of the reader during the first minute of reading. The order sequence was treated as a score (1 = 10 points, 2 = 9 points, etc.). Scores for each area were then added across the five readers (see Table 1). The time sequence order of attention is shown in Figure 5.

![Figure 4](image-url) Ten defined areas in the spread.
On average, the news about the carbon dioxide tax was noticed first. The positioning and design of that article helps to explain the result: this item is placed on a right-hand page, and right-hand pages often attract readers’ first glance (Wartenberg and Holmqvist, 2005). Also, this article has a large headline and is the dominant article on this page.

The article about the 39-year-old is in second place. Layout as well as position may explain this: it is the dominant text on the left-hand page and has a bold headline. As shown by Garcia and Stark (1991) and Wartenberg and Holmqvist (2005), the second article to be noticed is often on the left-hand page.

‘Valspalten’ (Election column) including ‘Frihetsfeminism’ (Freedom feminism) were in 3rd and 4th place of the text areas. This far right-hand column, perhaps conceived of as a whole, extends far up, which is typical of texts attracting readers’ attention at an early stage (cf. Wartenberg and Holmqvist, 2005).
Holmqvist, 2005). This column also has a picture, which makes it likely to receive early attention (Hansen, 1994; Wartenberg and Holmqvist, 2005).

'Studentilska' (Student anger) [PICTURE] and 'Studentilska' (Student anger) [TEXT] ended up in 5th and 6th place, respectively. Pictures usually attract attention before corresponding text and often function as entry points into texts. However, it is strange that the picture does not achieve a higher ranking as pictures are usually noticed early on in overall scanning. A plausible explanation for this deviation from the usual pattern is the picture's poor position, low on the right-hand page (Wartenberg and Holmqvist 2005). Also, the picture looks rather messy with no clear focal point.

Then the average reader's attention is drawn to the text area 'Två notiser' (Two news items). These news items are small and are placed outside the conventional column of news items. They could be thought of as insignificant, but are placed in a position which is looked at more than average (Hansen, 1998; Widman and Polansky, 1990; Wartenberg and Holmqvist, 2005).

Although it takes up the most space, 'Hela journalistpriset' (The entire journalist award) attracts attention quite late in spite of the fact that the headline VEM SKA HA PRISET? (Who will win the award?) is very large and on a red background, and that the readers in this study belong to the...
intended target group, i.e. newspaper people. However, none of these elements change the scores; maybe readers peripherally recognize this as an ad, which makes it score late in competing for reader attention (cf. Garcia and Stark, 1991; Wartenberg and Holmqvist 2005). A closer analysis shows that readers in fact read the text below the coloured headline before looking at the red text. ‘PIGGVAR’ (Turbot) ends up in last position because it is an ad in a fairly unfavourable position.

### 3.2.2 How long are different parts of the spread attended to?

The purpose of the second analysis was to map which areas in the spread were allocated the most time by the readers. For each reader, we calculated the time spent on each item of the newspaper spread as a percentage of the entire time spent on the spread. Then we calculated the average over different items (see Table 2). The average percentage time spent on the areas is illustrated in Figure 6.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Area</th>
<th>Average %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39-åringen (The 39-year-old)</td>
<td>39.3</td>
</tr>
<tr>
<td>2</td>
<td>Valspalten (Elections column)</td>
<td>16.5</td>
</tr>
<tr>
<td>3</td>
<td>Koldioxidskatten (The carbon dioxide tax)</td>
<td>10.9</td>
</tr>
<tr>
<td>4</td>
<td>Frihetsfeminism (Freedom feminism)</td>
<td>9.4</td>
</tr>
<tr>
<td>5</td>
<td>Två notiser (Two news items)</td>
<td>7.7</td>
</tr>
<tr>
<td>6</td>
<td>Studentilska (Student anger) [PHOTO]</td>
<td>6.8</td>
</tr>
<tr>
<td>7</td>
<td>Studentilska (Student anger) [TEXT]</td>
<td>4.7</td>
</tr>
<tr>
<td>8</td>
<td>Hela journalistpriset (The entire journalist award) [AD TEXT]</td>
<td>3.0</td>
</tr>
<tr>
<td>9</td>
<td>Piggvar (Turbot) [AD]</td>
<td>0.7</td>
</tr>
<tr>
<td>10</td>
<td>Vem ska ha priset? (Who will win the award?) [HEADLINE]</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 2 Average ranking of the different items in the newspaper spread with respect to the average percentage of time spent on each item
Results show that the ‘sex & crime’ text about ‘39-åringen’ (The 39-year-old) was read for the longest time (all 5 subjects read the article). The areas in 2nd and 4th place, respectively – the column ‘Valspalten’ (Election column) and ‘Frihetsfeminism’ (Freedom feminism) – were both up-to-date, provocatively formulated and well designed, and therefore absorbed a large part of the reading time. The area ‘Koldioxidskatten’ (Carbon dioxide tax) was in 3rd place as all the subjects read the text, although none of them read it for very long. Also the item in 5th place, ‘Två notiser’ (Two news items) attracted a fair amount of attention even though the two news items are short. The readers scanned ‘Studentilska’ (Student anger) [PHOTO] for a relatively long time. Pictures normally do not receive extensive attention, but this picture is messy and comprises texts and faces that require a longer viewing time. (Cf. Holsanova, 2001, in which the author studies picture viewing in detail by means of eye-tracking and connects her investigations to spoken picture descriptions.) ‘Studentilska’ [TEXT] was in fact read for a shorter time than the picture. ‘Hela journalistpriset’ is an ad with a generously spaced text, not taking long to read. However, only 3 out of 5 readers looked at it at all. The headline ‘VEM SKA HA PRISET?’ (Who will win the award?) was looked at by only 2 out of 5 readers and for a very short time. The ad ‘PIGGVAR’ (Turbot) was read least of all; only 2 out of 5 readers paid any attention to it.
In the results section of the second analysis, we observed that some items on the spread are generously spaced, while others contain many lines of text in a very restricted area. In order to obtain a fairer comparison, we finally calculated how closely the objects on the spread were read. The chosen measure for close reading or reading depth is milliseconds per square centimetres of each item. Using this measure, the following order was established.

By measuring reading time per area unit, the ranking order changes. ‘39-Åringen’ (The 39-year-old) is still leading with a reading attention of 207 ms/cm² on an average. This reading is about 5 times as intense as other larger texts, such as ‘Koldioxidskatten’ (The carbon dioxide tax). The text area ‘Två notiser’ (Two news items) moves up to 2nd place with a reading intensity of 134 ms/cm² on average. These results are in line with Hansen (1994), who also found that short texts are often read in more depth than longer ones.

### Table 3 The most carefully read items on the newspaper spread over the five readers

<table>
<thead>
<tr>
<th>Rank</th>
<th>Object</th>
<th>Milliseconds per cm² area of the objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39-åringen (The 39-year-old)</td>
<td>207</td>
</tr>
<tr>
<td>2</td>
<td>Två notiser (Two news items)</td>
<td>134</td>
</tr>
<tr>
<td>3</td>
<td>Frihetsfeminism (Freedom feminism)</td>
<td>110</td>
</tr>
<tr>
<td>4</td>
<td>Valspalten (Election column)</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>Studentilska (Student anger) [TEXT]</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>Studentilska (Student anger) [PHOTO]</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>Koldioxidskatten (The carbon dioxide tax)</td>
<td>40</td>
</tr>
<tr>
<td>8</td>
<td>Piggvar (Turbot) [AD]</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>Hela journalistpriset (The entire journalist award) [AD TEXT]</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>Vem ska ha priset (Who will win the award?) [AD]</td>
<td>1</td>
</tr>
</tbody>
</table>

3.2.3 How carefully were the different texts in the spread read?
In the results section of the second analysis, we observed that some items on the spread are generously spaced, while others contain many lines of text in a very restricted area. In order to obtain a fairer comparison, we finally calculated how closely the objects on the spread were read. The chosen measure for close reading or reading depth is milliseconds per square centimetres of each item. Using this measure, the following order was established.

By measuring reading time per area unit, the ranking order changes. ‘39-Åringen’ (The 39-year-old) is still leading with a reading attention of 207 ms/cm² on an average. This reading is about 5 times as intense as other larger texts, such as ‘Koldioxidskatten’ (The carbon dioxide tax). The text area ‘Två notiser’ (Two news items) moves up to 2nd place with a reading intensity of 134 ms/cm² on average. These results are in line with Hansen (1994), who also found that short texts are often read in more depth than longer ones.
Frihetsfeminism’ (Freedom feminism) ends up in 3rd place with 110 ms/cm², almost as intensely read as ‘Två notiser’ (Two news items). Other larger texts, i.e. ‘Valspalten’ (Election column) (48 ms/cm²), ‘Studentilska’ (Student anger) [TEXT] (45 ms/cm²), ‘Studentilska’ (Student anger) [PHOTO] (41 ms/cm²) and ‘Koldioxidskatten’ (The carbon dioxide tax) (40 ms/cm²) are all read with about the same intensity, considerably less than the top-scoring items. The least reading per area unit is found for the ads: ‘PIGGVAR’ (Turbot) (7 ms/cm²), ‘Hela journalistpriset’ (The entire journalist award) (6 ms/cm²) and ‘VEM SKA HA PRISET?’ (Who will win the award?) (1 ms/cm²). Reading of the ads amounts to about a 6th of the less-read larger texts, and to about a 30th of ‘39-åringen’ (The 39-year-old).

The result of the third step of analysis is illustrated in Figure 7.

**4. COMPARING GENERAL ASSUMPTIONS AND ACTUAL READING BEHAVIOUR**

The aim of this article was to track newspaper reading behaviour (especially entry points and reading paths) and to compare these to a semiotic analysis of newspaper layout and newspaper reading. In order to achieve this, we used the seven general assumptions about reading behaviour outlined in Kress and Van Leeuwen (1996). We now go through these general assumptions and compare them with the empirical results.
Generally, it should be noted that whereas eye-tracking data can answer the questions of in what order, for how long and how carefully the items on the newspaper spread were attended to, the assumptions extracted from these data can be interpreted only in terms of order. The semiotic framework, on the other hand, does not predict how much attention the items will be given nor how deeply they will be read.

Assumption 1 states that readers prefer new information and expect this to be on the right in the semiotic space. Applied to the newspaper spread it would mean, for example, that the article ‘Studentilska mot bostadsbrist’ (Student anger against housing shortage) would be read at an early stage. However, this hypothesis is not confirmed by the data from our experiment; ‘Studentilska’ was read relatively late, readers did not spend much time reading it, nor did they read it very carefully. Nevertheless, the right position can be advantageous for attracting early but very short attention as a result of the page-turning process (Wartenberg and Holmqvist, 2005).

Assumption 2 states that readers prefer the most general information at the top and the most specific information at the bottom of the semiotic space. Assumption 3 states that readers look for the most important information at the centre and less important information on the periphery. What these assumptions do not state is whether this also means that readers pay attention to items in this order, i.e. reading the top of the page first, then the bottom, and the centre of the page before the periphery. The top–bottom tendency is verified in our data. Significant results concerning the order of items attracting attention in the horizontal and vertical dimensions can be found elsewhere. For instance, Wartenberg and Holmqvist (2005) found that items at the top of the page were looked at significantly earlier than items lower down the page, and items on the left of a page were looked at significantly earlier than those on the right.

It has been assumed that readers look for graphically salient elements (assumption 4) and paratexts (assumption 5). Also the eye-tracking studies of newspaper design agree that pictures and headlines are the most common entry points. If we look at which items actually caught the eye of our readers during the first minute of reading the spread, we can see that all the readers looked at headlines, see Figure 8(a)–(e). Apart from the second reader in Figure 8(b), all subjects saw and processed the majority of the large headlines on the spread. The second reader had another strategy: his or her attention was caught by the headline ‘39-åringen’ (The 39-year-old) and he or she remained focused on it. This subject then decided to continue reading the article in depth during the rest of the time at the expense of further scanning. The next common entry point was pictures. We noted that 4 of the 5 readers saw the big picture ‘Studentilska’ (Student anger), even if the 5th reader in Figure 8(d) did so only very briefly. So far, we could say that assumptions 4 and 5 were confirmed by the eye-tracking data. One major exception, however, is the ad ‘VEM SKA HA PRISET?’ (Who will win the award?). Although colour is said to play an important role in drawing attention to an
Figure 8(a) The first minute of the first reader’s newspaper reading presented as scan paths.

Figure 8(b) The first minute of the second reader’s newspaper reading presented as scan paths.
Figure 8(c) The first minute of the third reader's newspaper reading presented as scan paths.

Figure 8(d) The first minute of the fourth reader's newspaper reading presented as scan paths.
entry point, it was not very attractive to our readers. Nor did the combination of size, font and layout help. Only 2 of the 5 readers fixed their gaze on the headline ‘VEM SKA HA PRISET?’ (Who will win the award?) and they did so very late.

Assumption 6 states that readers follow elements linked by framing devices such as lines and arrows. One example of this is the layout of the far-right column. In the scanning experiment, readers seemed to integrate some objects, such as the election column or the picture caption, with the corresponding article.

Assumption 7 states that readers scan the semiotic space before taking a closer look at certain units. In order to find out whether the eye-tracking data confirm this assumption, let us now take a closer look at Figures 8(a)–(e). The illustrations reveal five different reading behaviours that could be labelled in different ways.

Starting with the second reader, we notice that his or her strategy in Figure 8(b) differs most significantly from the other readers. This reader devotes his or her attention to only one article, ‘39-åringen’ (The 39-year-old), and is therefore labelled ‘the focused reader’. During the major part of the first minute, the focused reader systematically reads the headline and the text of this one article apart from a few very brief fixations on two articles that are spatially close.

The readers illustrated in Figures 8(c) and (d) could be called ‘editorial readers’ since they avoid the ads. Neither of them pays attention to the left part of the spread with ‘Hela journalistpriset’ (The entire journalist award) or the bottom ad ‘PIGGVAR’ (Turbot). But these two reading behaviours vary in several respects. First, they differ in their thematic choice.
of articles. While the first editorial reader illustrated in Figure 8(c) concentrates on ‘39-åringen’ (The 39-year-old) – apart from two small news items and headlines – the other editorial reader illustrated in Figure 8(d) focuses on ‘Frihetsfeminism’ (Freedom feminism), apart from one of the brief news items and headlines. Second, they vary in the number of paths through the spread as the first editorial reader illustrated in Figure 8(c) moves across the spread more frequently. Finally, we find differences in their way of scanning the picture: the first editorial reader studies the picture more carefully whereas the second one only gives it a brief glance.

We call the readers illustrated in Figures 8(a) and 8(e) ‘entry point overviewers’ since they visit the most important entry points such as headlines, coloured headlines, pictures, and the ad to the left. However, they show very little reading activity. Even here we find differences in their scanning and reading paths: the order of the visited entries differs and the reading activity is concentrated on other areas. An interesting point is that one of the entry point overviewers illustrated in Figure 8(e) is the only reader to connect the information from the photo belonging to ‘Studentilska’ (Student anger) with the contents in the headline and in the main text of the article. In sum, after having analysed all these different reading strategies, we can see that assumption 7 does not hold. Readers do not scan the semiotic space before taking a closer look at certain units.

5. DISCUSSION

Let us now comment on the fact that the five readers investigated form three groups: the entry point overviewers illustrated in Figures 8(a) and (e), the focused reader illustrated in Figure 8(b) and the editorial readers illustrated in Figures 8(c) and (d). Even though the number of subjects is small, it is possible to observe patterns tracing how their gaze moves across the entry points on the spread. The focused reader might seem to be an anomaly as he or she only reads one text on the spread, the text about a 39-year-old man prosecuted for child abuse. Of course, the designers of the spread do not want focused readers. How can they justify their work with an appetizing mixture of photos, paratexts and main texts if there are readers of this kind? How would those buying ads react if the focused reader were a common reader type? However, considering the endless possibilities of variation among the readers of a major newspaper like Svenska Dagbladet, the focused reader’s behaviour is not that strange. The focused reader probably has a special interest in this topic. What would the result be if one of the texts on the spread dealt with a topic which was of common interest to all the subjects in the test? To be more precise: what could be observed from a larger test group united by the same passionate interest in angling, for example, reading a spread with a news item about fishing?

The editorial readers are also dangerous readers from a newspaper editor’s perspective as they consistently avoid the ads, even the graphically
salient advertisement ‘Hela journalistpriset’ (The entire journalist award) ideally positioned to the left on the spread. The results of the eye-tracking analysis for this ad (especially its headline) stand in opposition to the semiotic analysis. The ad is read by only a few people, and when it is read, the specific information in the lower part is read before the general information in the headline at the top. A possible explanation for this is the common reluctance to pay attention to ads, a subject which is discussed in Rahm (in press), who concludes that interdiscursivity with non-commercial texts and genres is an important way of attracting readers’ attention in direct mail shots. This common reluctance to read ads can explain why readers in this study neglect the areas of ads even though ‘Hela journalistpriset’ (The entire journalist award) occupies a favourable position on the spread and is graphically salient.

The aim of this study was to track newspaper reading behaviour, especially entry points and reading paths, and to compare these with a semiotic analysis of newspaper layout and newspaper reading. We have presented convincing evidence that it is worthwhile testing hypotheses on entry points and reading paths by means of eye-tracking. It is hoped that this article makes a contribution to future interdisciplinary research on newspaper reading. Since this study contains eye-tracking data that can answer the questions of in what order, for how long and how carefully the items on the spread attracted attention, it offers a very nuanced analysis of reading behaviour.

There are many interesting topics to be studied in detail. For instance, how do readers perceive pictures, information graphics and captions that accompany a text (see Holsanova et al., forthcoming)? How do readers interpret the meaning of these different items and integrate them semantically? Holsanova (1999) gives an overview of different perspectives on text, pictures and their interplay, concluding with a call for further research, especially into the perception and interpretation of multimodality.

Semiotic or socio-semiotic analyses and eye-tracking studies can be mutually beneficial. If reading behaviour and reading paths are to be mapped in detail, semiotic analysis can be strengthened by being combined with the inductive approach of eye-tracking studies. Eye-tracking allows us to follow precisely what is being looked at and when. Gaze behaviour reveals the paths through the text, paths through the picture and the connections created between picture and text. When eye-tracking data are supplemented with simultaneous or retrospective verbal protocols and interview methods (see Holsanova, 2001; Holsanova and Holmqvist, 2004), this results in a very detailed view of reader interaction with newspapers. We can then investigate reader behaviour, the rationality behind that behaviour, as well as reader expectations and attitudes.

Conversely, detailed data from eye-tracking and verbal protocols benefit from socio-semiotic analyses, where a deductive approach contributes to understanding. The semiotic space is contextualized at different producer
and consumer levels, and has an inherent socio-cultural frame. The analytical tools used (dimensions of information structure, paratext, etc.) provide another way of analysing the material. In conclusion, the combination of the inductive approach of the eye-tracking analysis and the deductive approach of the semiotic analysis offers rich possibilities for the development of challenging interdisciplinary research in the field of newspaper reading.

NOTES
1. It should be noted that each reader might give priority to one or a few of these seven principles when creating meaning of the semiotic space.
2. These studies include: Küpper (1989), Widman and Polansky (1990), Garcia and Stark (1991), Hansen (1994), Lewenstein et al. (2000), Lundqvist and Holmqvist (2001), Barthelson (2002), Holmqvist et al. (2003), Holsanova et al. (2003), Holsanova and Holmqvist (2004), Wartenberg and Holmqvist (2005), Holmqvist and Wartenberg (2005), Holsanova et al. (forthcoming) as well as this research.
3. The eye-tracker used is SMI iView RED II 50 Hz pupil/corneal reflex.
4. See Holmqvist et al. (2003) on the proportion of scanning and reading in newspapers on the web as well as newspapers in traditional paper format.
5. Assuming a given line distance and font size, this measure directly translates to milliseconds per letter. Hansen (1994) uses another measure of reading depth: the reading depth of an article is defined as the percentage of text being read in relation to all the text in an article.

ACKNOWLEDGEMENTS
The work described in this article was partly supported by a Fellowship Grant to Jana Holsanova from the Swedish Research Council (Vetenskapsrådet).

REFERENCES
Hansen, J.P. (1994) 'Analyse af Læsernes Informationsprioritering (Analysis
of readers’ information priorities), Kognitiv Systemgruppen Forskningscenter Risø, Roskilde, July.


**Biographical Notes**

JANA HOLSANOVA is a Post-Doctoral Research Fellow at the Cognitive Science Department at Lund University, Sweden. She worked on a number of international and interdisciplinary research projects dealing with identity formation, positioning, storytelling and intertextuality in dialogues. Her current project concerns cognitive aspects of language use and the integration of visual perception and language production. An important topic in her publications is multimodal discourse and the interplay between language and pictures.

Address: Lund University, Kungshuset, Ludagård, 222 22 Lund, Sweden. [email: jana.holsanova@lucs.lu.se]
HENRIK RAHM is a lecturer in linguistics at the University College of Kristianstad where he teaches courses in grammar, phonetics, semantics, discourse analysis and non-fiction texts. He also works at the Department of Scandinavian Languages at Lund University, amongst other things on the EU research project PARADYS, a cross-disciplinary project 2001–4 involving linguists and sociologists in seven European countries dealing with social positions and citizen participation in GMO discourse. Other interests are genre analysis, text patterns and semiotics in newspapers as well as the interdiscursivity of advertisements.

Address: Department of Humanities and Social Science, University College of Kristianstad, SE-291 88 Kristianstad, Sweden [email: henrik.rahm@husa.hkr.se] Department of Scandinavian Languages, Lund University, Box 201, SE – 221 00 Lund, Sweden. [email: henrik.rahm@nordlund.lu.se]

KENNETH HOLMqvist is an Associate Professor at the Cognitive Science Department at Lund University, Sweden. He is now the head of the Eye-Tracking Laboratory at Lund University, using eye-tracking data to assess cognitive processes. He has published over 50 papers in areas of applied eye-tracking and visual perception.

Address: Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden. [email: kenneth.holmqvist@lucs.lu.se]