Discourse-based lexical anticipation : the nature and contextual basis of predictions in language comprehension

Otten, M.

Publication date
2008

Chapter 5

Does appearance matter? A self-paced reading study exploring whether linguistic predictions extend to the visual form.

Sometimes a sentence or story is so constraining that we already know what our conversation partner is going to say, or how the paragraph will end, before we have even heard or read the entire utterance. In this study we have explored whether these linguistic predictions are limited to semantic and lexical features, or whether they also include the visual form of the expected word. In a self-paced reading experiment, participants were shown sentences that contained expected, unexpected or neutral nouns. These nouns could be presented in the standard font (same font as the other words in the story) or in a deviant font. We argued that, if people actually predict the visual form of an expected word, then an expected word in a deviant would mismatch this word-image prediction. We thus expected that the slowing effect of reading a deviant font word would be larger for expected words in a predictive context than for the neutral context. As expected, we observed that readers slow down more when they encountered a deviant typeface noun in a constraining discourse than in a neutral discourse. This suggests that readers make predictions about the actual visual appearance of the upcoming word. However, unexpected nouns showed a similar font-effect as the expected nouns, which could be taken to suggest that other factors than word-image prediction play a role. Thus, even though the findings indicate that linguistic predictions could have a visual component, this experiment does not provide the definitive answer.

Introduction

Electrophysiological studies suggest that people can use the linguistic context to make an informed guess about the word that will follow (DeLong et al., 2005; Otten et al., 2007; Otten & Van Berkum, in press, Van Berkum, 2005 #452; Wicha, Bates et al., 2003; Wicha, Moreno et al., 2003; Wicha et al., 2004). Levelt, Roelofs and Meyer (1999) suggest that the lexical representation of each word consists of three different information-stores: the lexical concept of the word, the lemma (which contains syntactic information about the word), and the lexeme/ morpheme (which specifies the phonological and the orthographic contents of the word). From the above-mentioned ERP studies we know that linguistic predictions include not only information about the meaning of the upcoming word, but also information about the syntactic properties of the word (in this case the grammatical gender Otten et al., in press; Van Berkum et al., 2005; Wicha et al., 2004). This shows that, when we predict a word, the lemma that is related to the anticipated word is also pre-activated. A recent study by DeLong, Urbach and Kutas (2005) suggests that linguistic predictions can also include the phonological information included in the lexeme. In this study Delong et al. used the fact that in English articles vary depending on the initial sound of the following word ('a' before words beginning with a consonant and 'an' before words beginning with a vowel). They observed a differential ERP to articles that were inconsistent with the initial sound of the predicted word. From this DeLong and colleagues conclude that linguistic predictions extend to the lexeme.

In this experiment we have explored whether linguistic predictions can go even beyond the lexeme, predicting exactly what the visual form of the predicted word will be, on paper or on the screen. To make predictions about the visual form of the word, first the orthographic structure, i.e. the letters of which it is composed, needs to be retrieved from memory. However, since the actual visual manifestation of these letters is not static (i.e. compare the visual form of "a" and "A"), pre-activation of orthographic information alone is not enough to make predictions about the visual form of a word. For this, readers need to combine memorised orthographic information with knowledge about the standard font in which the letters are presented. In this experiment, we have thus tested whether people are capable to translate their specific lexical prediction into a visual prediction, by combining orthographic information with dynamic, context-dependent information about the standard typeface.

Research on visual perception indicates that predictions can influence visual processing. Objects are more easily recognised when they are seen within, or shortly after, a consistent setting (i.e. a pan in a kitchen setting

Biederman, Mezzanotte, & Rabinowitz, 1982; Davenport & Potter, 2004; Palmer, 1975). For simple visual stimuli, there is evidence that predictive processes actually modify what we perceive (Jolij & Lamme, submitted). Taken together, these findings could be taken to suggest that anticipatory processes can influence visual perception (cf. Bar, 2004). When we combine these findings with the evidence that readers and listeners make highly specific predictions about upcoming words, this suggests that predictions about upcoming linguistic content could also include very specific predictions about the visual word-image.

We have tested this hypothesis in a self-paced reading experiment, which allows us to monitor the time it takes a reader to process each word in a sentence. We created stories that induced specific lexical predictions, such as Example story 1 in Table 5.1. In this story, the word *apotheek* (*pharmacy*) is highly expected to appear. Based on previous experiments (Otten et al., 2007; Otten & Van Berkum, in press) we know that readers use such a predictive context to pre-activate the supported word at the lexical level (lexical prediction). By combining knowledge about the orthography of the predicted word (i.e. the representation of the word in terms of its successive graphemes, i.e. "a" followed by "p" followed by "o" etc), and knowledge about the specific form of the letters in this context (i.e. the standard font) the reader can create a very detailed prediction about the visual word-image of this lexical prediction, i.e. apotheek.

If a reader indeed creates a word-image prediction, then seeing the expected word in an unexpected font ("apotheek") will mismatch the predicted visual word-image. Such a word-image mismatch will then lead to slower reading times. Of course, any differences in reading time between the word-image prediction match "apotheek" and the word-image prediction mismatch "apotheek" can not be solely attributed to the word-image mismatch, because besides a specific word-image mismatch, these nouns also have an unexpected and deviating font relative to the preceding context. To measure the more general effect that such an unexpected font has on the reading time of a word we also created non-predictive stories (Example 2 in Table 5.1). In these stories, the discourse did not support a lexical prediction. Any differences in reading time between "apotheek" and "apotheek" will thus be elicited by the difference in physical deviance for the noun in the standard and the deviating font.

Table 5.1. An example of the items used in experiment 1A, in the original Dutch version and an approximate English translation. The critical word is underlined and shown in the standard font and the deviating font.

Example 1: Expected Noun

Na het bezoek aan de streekdokter moest de boerin nog medicijnen ophalen. Ze ging daarvoor naar een apotheek/ apotheek in de grote stad.

After visiting the local doctor the farmer's wife still had to pick up the medication. For that, she went to the pharmacy/ pharmacy in the city.

Example 2: Neutral Noun

Piet fietste naar een apotheek/ apotheek om zijn medicijnen op te halen. Omdat hij daar toch was nam hij direct de pil voor zijn vriendin mee.

Piet cycled over to the pharmacy/ pharmacy to pick up his medication. On behalf of his girlfriend he also collected a prescription for the pill.

Example 3: Unexpected Noun

Na het bezoek aan de streekdokter moest de boerin nog medicijnen ophalen. Ze ging daarvoor naar een ziekenhuis/ ziekenhuis in de grote stad.

After visiting the local doctor the farmer's wife still had to pick up the medication. For that, she went to the hospital/ hospital in the city.

In the predictive story "apotheek" is presented in a deviating font *and* violates the word-image prediction, whereas in the neutral story, "apotheek" only has a deviating font. The critical comparison to test for the presence of a specific word-image prediction is thus between the predictive and the neutral stories. Both neutral and unexpected nouns presented in a deviant font will show *font-costs*, i.e. both types of words will be read slower when presented in a deviating font than in the standard font. However, if readers use a constraining discourse to predict the word-image, then these costs will be higher for the expected words than for the neutral words. The font-costs for neutral deviant font words reflect only the increase in reading times that is the consequence of simply reading a word in a font that differs from the standard font. In a constraining context, however, a reader can use the context to create a word-image prediction. If the expected word is then presented in a deviant font, the perceived word-image will diverge from the predicted word-image. In this case, the font-costs for the expected deviant font word will reflect the basic 'deviant font'-costs with the added 'unexpected word-image'-costs, and will thus be higher than for the neutral deviant font words. If readers do not make predictions about the specific word-image, based on the constraining context, then the font-costs for the neutral and expected words will be comparable.

To explore the interaction between meaning mismatches and the font manipulation, we also included constraining stories in which the expected word was replaced by a coherent but unexpected noun (see Example 3, Table 5.1). If a reader has a strong prediction for a specific target word (*apotheek*), but reads another, less expected noun ("ziekenhuis"), then both the lexical prediction and the word-image prediction will be violated at the same time. This double violation exists when the unexpected noun is written in the standard *and* in the deviant font. Unexpected nouns will be read slower than the expected nouns when presented in standard font, because unexpected nouns violate the prediction at a lexical and visual level. There will also be font-costs for the unexpected noun when it is presented in a deviant font, because of the general slowing that is the consequence of reading words in an unexpected font. If the font-costs for the unexpected word indeed only represent the basic 'deviant font'-costs, the font-costs for the unexpected words should be comparable to the neutral words.

Methods

Participants

57 students of the University of Amsterdam participated for course credit. 1 participant was excluded from the analysis because of technical problems during the experimental session, 7 others were excluded because of low scores on questions about the content of the sentences. The 49 remaining participants (42 females; mean age 20 years, range 18-33) were all native speakers of Dutch.

Materials

The critical materials for this experiment consisted of 180 two-sentence stories. Each story was created to be constraining, so that people predicted one specific word (cloze values: range from 73% to 100%, mean = 86%, sd = 8%). If the story actually contained the predicted word, the noun matched the prediction of the reader at the level of meaning (independent of the actual font that word is presented in). The predictive stories could also contain another noun than the predicted noun, which was coherent but less expected (cloze range 0% to 20%, mean = 1% ,sd = 3%). These unexpected nouns mismatch the prediction at the lexical level and at the level of the visual word-image (independent of the actual font that word is presented in). For each expected target word we also designed a neutral context story. In these neutral stories the critical noun always appeared very early in the first sentence of the story, such that the reader had not formed any specific prediction. These neutral nouns are never inconsistent with a lexical or word-image prediction, since the context does not give rise to any prediction. In all experimental conditions the critical noun could be shown in the same font as the rest of the story (Courier New), or in a deviating font (Schule 1995, resembling handwriting as it is taught in primary school). For the expected words, presented in a predictive context, the deviant font creates an additional word-image mismatch, which is not present in the neutral and unexpected deviant font words.

We constructed 60 filler items that resembled the structure of the critical items, thus also consisting of two sentences. One word in the first sentence of a filler item was always shown in the deviating Schule font. This deviating word was never a noun, or the first or last word. Each filler item was accompanied by a question about that story. The filler items were randomly mixed with the 180 experimental items. Of the experimental items, 60 items contained an expected noun, 60 items contained an unexpected noun and 60 contained a neutral noun. For each of these three types of stimuli the critical noun was either shown in the standard courier font or in the deviating Schule font, thus resulting in 30 items for each of the six conditions (expected standard font, expected deviant font, unexpected standard font, unexpected deviant font, neutral expected font, neutral deviant font). By rotating the conditions in this list, 6 more lists of stimuli were created. Each of the 6 lists contained all 180 experimental stimuli, and 60 identical filler items. Each participant was shown one of these six lists of stimuli, so that one participant saw all the stimuli, but never in more than one condition.

Procedure

We presented the stories in a standard noncumulative moving-window self-paced reading paradigm. Subjects read through each story word by word, with each button press disclosing the next word while replacing all other letters in the story with hyphens. As they pressed their way through a story, subjects could see its overall sentential and formatting layout (including punctuation) as well as the position of the currently visible word therein. Subjects were asked to process each story for comprehension and to adapt their speed to this. Simple yes–no comprehension questions were asked after the filler items, which made up 25% of the stories, to keep the participants focused on the content of each story. A reading session consisted of five trial blocks separated by a short break, and took approximately 60 minutes on average.

Analysis

Participants that answered more than 10% of the questions incorrectly were excluded from the analysis. Based on this criterion 7 participants (mean percentage correct: 66%, sd = 40%) were excluded. The remaining 49 participants answered 94% of the questions correctly (sd = 2.9%). Two items were not shown correctly during the experiment, and were thus excluded from the analysis, which as a result included 178 items.

We analyzed reading times in two regions of each story, namely a baseline-region that consisted of the three words that preceded the critical noun, and an experimental region that included the critical noun and the four words that followed this noun. Before analysis, outlying reading times were eliminated. Reaction times that deviated more than 2 standard deviations from the mean reading time for the subject in that condition *and* from the mean reading time for the item in that condition were removed. As a result 1.5% of the data, evenly distributed across the 8 X 6 cells of the design (ranging from .7% to 2.2%), was excluded from the analysis.

For each of the 8 wordpositions in the critical regions, we computed mean reading times per subject and per item for each of six conditions. For each word position we examined the resulting reading times in a by-subjects and a by-items two-way ANOVA with a factor Predictability (three levels: Expected Noun, Neutral Noun and Unexpected Noun) and a factor Font Consistency (two levels: Standard Font and Deviant Font). Conservative degrees of freedom were employed when violations of sphericity were found(epsilon <1). The Huynh–Feldt correction was used for smaller violations of sphericity (values between 0.75 and 1.00) while the Greenhouse–Geisser correction was employed for more severe violations of sphericity (epsilon <0.75). Significant results that concerned the factor Predictability where followed by planned comparisons to establish which levels of this factor differed.

Results

Figure 5.1 shows the reading times for all words in the baseline and critical region (for the exact reading times see Table 5.2). The bars in Figure 5.2 represent the corresponding font-costs for the critical region only. As expected, it is clearly visible that all deviant font nouns (grey lines at position 0) elicit longer reading times. This effect persists in the words that follow the critical nouns even though these words are all presented in standard font. Significant main effects of Font Consistency at the critical noun and the two words that follow (CW, CW 1 and CW 2, see Table 5.3) corroborate this observation. At the position of the critical word itself, the effect of a deviant font is significantly modulated by the predictability of the critical noun (Table 5.3, significant interaction between Predictability and Font consistency).

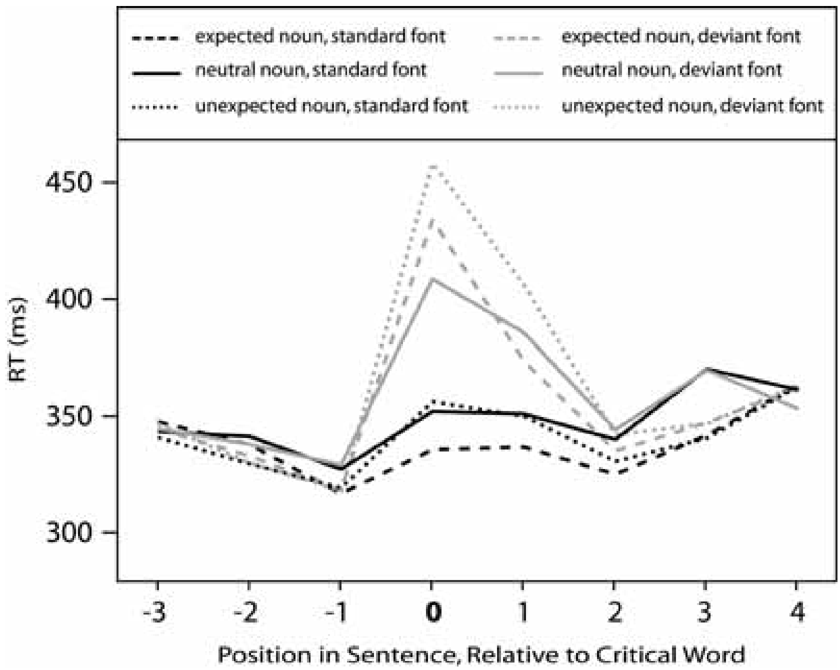

Figure 5.1 *Average reading times for all six experimental conditions, plotted for the 8 consecutive wordpositions that make up the baseline and critical region. The critical noun itself is referred to as position 0.*

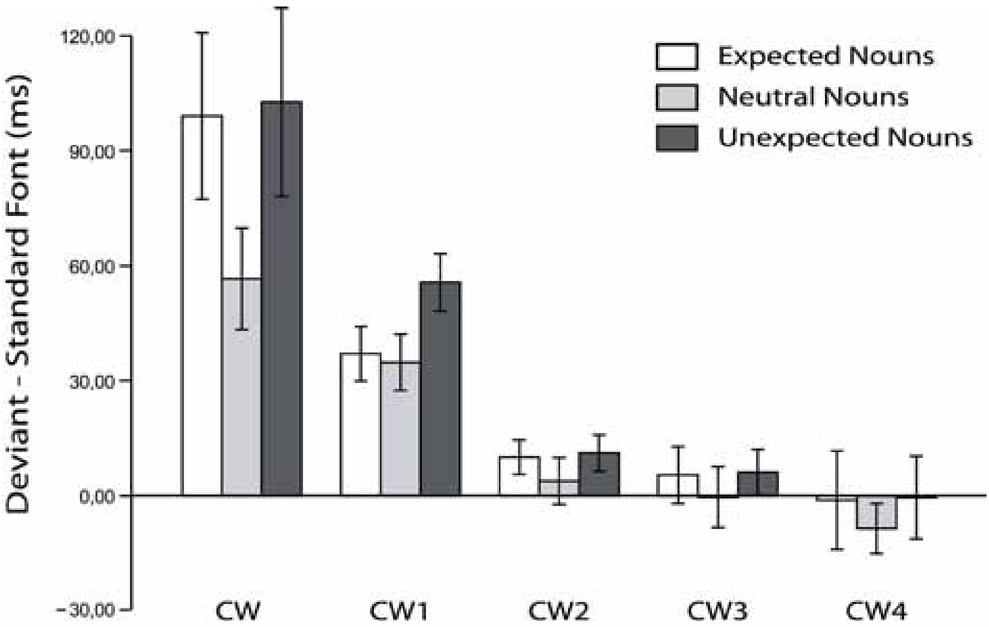

Figure 5.2 *The effect of Font Deviance (the difference in reading-time between nouns in a deviant and a standard font), depicted for the 3 different types of nouns. Error bars indicate the Standard Error.*

Table 5.2. Mean reading times for the words in the Baseline and Critical region.

		CW-3	CW-2	CW-1	**CW**	CW 1	CW 2	CW 3	CW 4
Expected Noun	mean	347	337	316	335	336	324	341	363
Standard Font	*SD*	*149*	*144*	*105*	*166*	*130*	*117*	*184*	*225*
Expected Noun	mean	344	332	317	433	373	334	346	361
Deviant Font	*SD*	*130*	*129*	*112*	*430*	*172*	*122*	*185*	*229*
Neutral Noun	mean	343	341	327	352	350	340	370	361
Standard Font	*SD*	*138*	*142*	*109*	*185*	*151*	*139*	*214*	*173*
Neutral Noun	mean	344	337	329	408	385	344	369	353
Deviant Font	*SD*	*152*	*144*	*131*	*325*	*175*	*153*	*263*	*168*
Unexpected Noun	mean	340	329	318	355	349	330	340	362
Standard Font	*SD*	*137*	*120*	*108*	*200*	*129*	*115*	*143*	*214*
Unexpected Noun	mean	346	329	318	458	406	342	346	363
Deviant Font	*SD*	*171*	*122*	*109*	*451*	*189*	*121*	*151*	*229*

Table 5.3. Overview of the ANOVAs for the critical region and the baseline region. Reported here are the results of a by subjects analysis (F_1, ε_1 and p_1) and a by items analysis (F_2, ε_2 and p_2).

		df	**CW –3**	**CW -2**	**CW -1**	**CW**
Predictability of Noun	F_1	*2,96*	.08	1.59	4.83	3.44
	F_2	*2,354*	.18	1.23	9.15	4.97
	ε_1		.82	-	.82	.77
	ε_2		-	.88	.95	.97
	p_1		.89	.20	.01	.05
	p_2		.83	.29	<.001	<.01
Font Consistency	F_1	*1,48*	.07	2.55	01	25.75
	F_2	*1,177*	.04	1.16	.02	71.23
	p_1		.79	.12	.95	<.001
	p_2		.85	.28	.89	<.001
Predictability * Font	F_1	*2,96*	.55	.30	.08	3.39
	F_2	*2,354*	.82	.27	.01	3.19
	ε_1		-	-	-	-
	ε_2		.94	-	.97	.95
	p_1		.58	.75	.93	.04
	p_2		.44	.76	.99	.05
		df	**CW 1**	**CW 2**	**CW 3**	**CW 4**
Predictability of Noun	F_1	*2,96*	9.89	4.50	11.14	.47
	F_2	*2,354*	11.60	3.97	10.00	.76
	ε_1		-	-	-	-
	ε_2		-	.87	.70	.62
	p_1		<.001	.01	<.001	.63
	p_2		<.001	.03	<.001	.41
Font Consistency	F_1	*1,48*	74.84	10.80	.87	.30
	F_2	*1,177*	179.6	11.49	1.53	.01
	p_1		<.001	<.01	.35	.58
	p_2		<.001	.001	.21	.92
Predictability * Font	F_1	*2,96*	2.98	.51	24	.19
	F_2	*2,354*	3.93	.84	.18	.64
	ε_1^a		.92	-	-	.89
	ε_2^a		-	.98	.88	.86
	p_1		.06	.60	.79	.83
	p_2		.02	.43	.81	.50

Table 5.4. Planned comparisons for the interaction between Predictability and Font Consistency				
		df	**CW**	**CW 1**
Predictability (Expected vs. Neutral)	F_1	*1,48*	4.42	.08
* Font Consistency	F_2	*1,177*	4.25	.01
	p_1		.04	.79
	p_2		.04	.98
Predictability (Unexpected vs. Neutral)	F_1	*1,48*	5.62	5.56
* Font Consistency	F_2	*1,177*	5.67	6.25
	p_1		.02	.02
	p_2		.02	.01
Predictability (Expected vs. Unexpected)	F_1	*1,48*	.04	2.92
* Font Consistency	F_2	*1,177*	.33	5.46
	p_1		.85	.09
	p_2		.56	.02

Planned comparisons (Table 5.4) show that the increase in reading times that accompanies a font mismatch is much larger for the expected than for the neutral nouns. Surprisingly, the increase in reading times for a deviating noun is also larger for unexpected than for neutral words.

The word that follows the critical noun (CW 1, which is always presented in standard font) still shows an increase in reading times as a result of the deviating font (Table 5.3, main effect of font consistency). At this position, the increase in reading times is roughly equal for the expected and the neutral noun, whereas the unexpected nouns show a larger increase than the nouns in the other two conditions (Table 5.4). Figure 5.2 clearly illustrates this difference between the critical word (CW) and the word that follows (CW 1) with regard to the interaction between noun predictability and font consistency.

The results in Table 5.3 also show a main effect for the noun predictability, which lasts from 1 word before the critical noun until 3 words after the critical noun. Planned comparisons (Table 5.5) show that the effect at CW -1 follows from a difference between neutral stories on the one hand, and predictive stories (with expected and unexpected words) on the other hand, with neutral trials resulting in slower reading times. These slower reading times for neutral stories are also visible at CW 2 and CW 3. At CW (the critical

Table 5.5. Planned comparisons for the main effect of Predictability							
		df	**CW -1**	**CW**	**CW 1**	**CW 2**	**CW 3**
Expected vs. Neutral	F_1	*1,48*	6.04	2.58	6.36	7.73	12.73
	F_2	*1,177*	14.14	.06	7.61	6.18	12.78
	p_1		.18	.61	.02	<.01	.001
	p_2		<.001	.82	<.01	.01	<.001
Unexpected vs. Neutral	F_1	*1,48*	5.51	3.87	3.72	1.86	19.05
	F_2	*1,177*	9.55	6.40	3.92	.43	9.32
	p_1		.02	.06	.06	.18	<.001
	p_2		<.01	.01	.05	.51	<.01
Expected vs. Unexpected	F_1	*1,48*	.08	6.49	19.33	3.44	.024
	F_2	*1,177*	.87	7.03	24.90	6.58	1.37
	p_1		.78	.01	<.001	.07	.88
	p_2		.35	<.01	<.001	.01	.24

noun) and CW 1, expected words are read at a faster pace than unexpected and neutral words.

Discussion

Our results show that for expected words the font-costs (i.e. the difference in reading times between words in a standard and a deviating font) are higher than for neutral words. This pattern of results suggests that readers make very specific predictions about the visual manifestation of highly predictable words. It thus seems that when readers make specific linguistic predictions, they also pre-activate the orthographic information associated with the predicted word. Perhaps even more interestingly, these results also suggest that readers are able to quickly integrate the activated orthographic information with the relevant contextual information about the standard font, to reach a specific visual prediction.

The fact that people are able to use contextual information to anticipate upcoming words not only at the lemma level (Otten et al., 2007; Otten & Van Berkum, in press; Van Berkum et al., 2005; Wicha, Bates et al., 2003; Wicha, Moreno et al., 2003; Wicha et al., 2004) but also at the level of the lexeme (DeLong et al., 2005, this experiment) is consistent with an interactive connectionist model of language perception as postulated by McClelland and Rummelhart (McClelland & Rumelhart, 1981). In this model activation of a word at the lexical level will automatically activate the related letters, and this activation consequently extends to the visual features that make up these letters. Although in the original model activation at the word level comes from real visual (or auditory) input, this model also acknowledges the role of higher level, top down input.

The difference in font-costs at the expected and neutral nouns thus suggests that linguistic predictions could extend to the visual features of the expected word. Our results, however, show one surprising finding. Unexpected words show almost identical font-costs as the expected words. However, according to our initial hypotheses, the font-costs for the unexpected words should have been comparable to the neutral words. This unexpected finding raises the possibility that the font-costs for the expected and unexpected words both have the same origin, which does not lie in visual linguistic predictions. In constraining stories (which contained the expected and unexpected words) the target noun was presented in the second sentence, whereas in the neutral stories the target word was always presented very early in the first sentence. The neutral stories thus differ from the predictive stories in the position of the critical noun. This leads to an overall difference in reading-times between the neutral and predictive stories: in the critical region, the predictive stories were read slower than neutral stories.

But can these overall differences in reading speed due to critical word position explain the larger font-costs for expected and unexpected words? Reading speed is higher for the constraining context than for the neutral context. One could thus reason that spillover-effects are more likely for deviant-font words that are presented in a constraining context: readers are pressing the button at a relatively fast rate, which increases the chances of a delayed response to the deviant noun (i.e. an increase in reading times that only becomes clear at one of the words that follows the critical word). If anything, however, a 'fixed fast rate' confound should lead to *smaller* font-costs for critical nouns that are presented later in the story, not *larger* font-costs.

Furthermore, the words that follow the critical nouns show that the processing of deviant expected and deviant unexpected words is not identical. The font-costs for unexpected words remain higher than for neutral words also after the critical word has been presented, whereas for the expected word this effect equals out immediately after reading the critical word. If the position of the critical word within the story is the only determiner of the difference between predictive and neutral stories, then the effects of expected and unexpected words should be the same, because these two words have exactly the same position in exactly the same story.

Taken together, it does not seem likely that the difference in word position is in itself responsible for the differences in font-costs between expected and unexpected nouns on the one hand and neutral nouns on the other hand. This would suggest that the difference in font-costs between neutral and expected words can indeed be attributed to the presence of a visual word-image prediction in the latter. But how then, if not by contextual factors, can we explain the additional processing costs that are visible for the unexpected deviant font words? Perhaps this answer lies in the extra emphasis that is indicated by a deviant font (McAteer, 1992). In contrast to the neutral and expected nouns, unexpected nouns were not consistent with expectations. Sanford and colleagues (2006) observed that when a word was foregrounding by a deviating typeface inconsistencies were noted more often than in sentences that did not contain such visual emphasis. It is thus possible that the extra emphasis on the mismatch, due to the conspicuous font, has made the meaning mismatch more striking, which in turn has led to the increased times for the unexpected nouns in comparison with the neutral nouns.

Conclusion

The results are consistent with the idea that people make predictions about the visual form of the expected word. The current experiment, however, does not allow us to conclude with certainty that the effects observed for neutral and expected words can be traced back exclusively to predictive processes. We are thus currently replicating this experiment with a neutral condition that resembles the predictive condition in length and critical word position. This design will allow us to determine whether the differences between the font effect for expected and neutral words can be attributed to the difference in critical word position, or whether this difference is due to the visual component of the linguistic prediction.