Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 5, 862–877

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.5.862

Effects of Contextual Predictability and Transitional Probability on Eye Movements During Reading

Steven Frisson
New York University and University of Massachusetts at Amherst

Keith Rayner
University of Massachusetts at Amherst

Martin J. Pickering
University of Edinburgh

In 2 eye-movement experiments, the authors tested whether transitional probability (the statistical likelihood that a word precedes or follows another word) affects reading times and whether this occurs independently from contextual predictability effects. Experiment 1 showed early effects of predictability, replicating S. A. McDonald and R. C. Shillcock's (2003a) finding that words with a high transitional probability (*defeat* following *accept*) are read faster than words with a low transitional probability (*losses* following *accept*). However, further analyses suggested that the transitional probability effect was likely due to differences in predictability rather than transitional probability. Experiment 2, using a better controlled set of items, again showed an effect of predictability, but no effect of transitional probability. The authors conclude that effects of transitional probability are part of regular predictability effects. Their data also show that predictability effects are detectable very early in the eye-movement record and between contexts that are weakly constraining.

Keywords: eye-tracking, transitional probability, contextual predictability

The predictability of a word in a context is known to affect reading times. Recently, McDonald and Shillcock (2003a, 2003b) argued for another type of predictability effect, *transitional probability* or the statistical likelihood that a word will precede or follow another word. They argued that transitional probability operates independently from "regular" predictability effects, at least when predictability is low. In this article, we examine whether transitional probability effects are truly independent from predictability effects and, if so, whether they can be overridden in more constraining contexts.

A considerable amount of research has shown that the *predictability* of a word in a context influences the time needed to process that word. Generally, the more predictable a word is given the context, the faster or easier that word is processed. This finding has been shown in naming (McClelland & O'Regan, 1981), lexical decision (Fischler & Bloom, 1979; Schwanenflugel & LaCount, 1988; Schwanenflugel & Shoben, 1985), speech monitoring (Cole & Perfetti, 1980), event-related potentials (e.g., Kutas, Lindamood, & Hilliard, 1984), and a large number of eye-movement studies (e.g., Ehrlich & Rayner, 1981; Rayner & Well, 1996). We concentrate on the last kind of study.

A word is highly predictable when the likelihood that that word will follow a given context fragment is high. Thus, predictability is often defined in terms of "contextual constraint." When a context is highly constraining, that is, when a target word is highly constrained, only a very limited number of words are likely to complete the context fragment; when the context is low constraining, many words will fit the fragment. For example, *stamp* is highly predictable in the (high-constraining) context, *He mailed the letter without a* ___, but low predictable in the (low-constraining) context *He saw a beautiful* ___. Estimates of predictability are generally compiled by the use of a modified Cloze task, in which participants are asked to complete sentence fragments, and the probability of a certain word appearing in a certain fragment is calculated.

Predictability has been shown to influence a variety of eye-movement characteristics. While there are some slight differences between experiments regarding when and how predictability effects show up, there is convincing evidence that high-predictable words are read faster than low-predictable words, with this difference appearing on the first fixation on the word (Altarriba, Kroll, Sholl, & Rayner, 1996; Binder, Pollatsek, & Rayner, 1999; Rayner, Ashby, Pollatsek, & Reichle, 2004), on single-fixation

Steven Frisson, Department of Psychology, New York University, and Department of Psychology, University of Massachusetts at Amherst; Keith Rayner, Department of Psychology, University of Massachusetts at Amherst; Martin J. Pickering, School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh, Scotland.

This research was conducted while Steven Frisson was at the University of Massachusetts funded by a postdoctoral fellowship from the Belgian Science Foundation. The research was also supported by National Science Foundation Grant 0236732 (awarded to Brian McElree) and Grants HD17426 and HD27650 from the National Institutes of Health to Keith Rayner.

Portions of these data were presented at the City University of New York conference, March 2004, College Park, Maryland, and at the Architectures and Mechanisms for Language Processing conference, September 2004, Aix-en-Provence, France. We thank Maryellen MacDonald, Richard Shillcock, and Barbara Juhasz for their helpful comments.

Correspondence concerning this article should be addressed to Steven Frisson, who is now at the School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom.

duration (Lavigne, Vitu, & d'Ydewalle, 2000), on gaze duration (the sum of all fixations on a word before fixating another word; Balota, Pollatsek, & Rayner, 1985; Binder et al., 1999; Lavigne et al., 2000; Rayner, Ashby, et al., 2004; Rayner & Well, 1996), on rereading time (Calvo & Meseguer, 2002), and on total time spent reading the target word (Calvo & Meseguer, 2002; Ehrlich & Rayner, 1981; Rayner & Well, 1996). There is also evidence that predictability has an effect on "spillover time," the first-fixation duration following the target word (Balota et al., 1985; Calvo & Meseguer, 2002). More predictable words are also skipped more often than less predictable words (Altarriba et al., 1996; Balota et al., 1985; Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Ehrlich & Rayner, 1981; Rayner, Ashby, et al., 2004; Rayner & Well, 1996) and have a lower probability of being refixated (Calvo & Meseguer, 2002; Ehrlich & Rayner, 1981; Inhoff, 1984).

Schwanenflugel and LaCount (1988; see also Schwanenflugel & Shoben, 1985; Stanovich & West, 1979, 1981) hypothesized that a higher number of featural restrictions can be generated from high-constraining contexts, making the set of possible completions smaller, thereby raising readers' expectations that a particular word will follow. However, the influence of predictability *strength*, or the degree of contextual constraint, is not very well understood. Typically, experiments are set up using target words that differ substantially in Cloze value, often with probabilities of .70 to .90 for high-predictable words and less than .10 for low-predictable words. According to Hyönä (1993), who did not find predictability effects when comparing conditions with Cloze probabilities of .65 and .32, strong contextual manipulations are necessary in order to find predictability effects on early reading times. However, further research has shown that rather than a simple difference in contextual strength, it also depends where on the scale this difference is taken from. Rayner and Well (1996) showed that gaze durations on high- (.86) and medium- (.41) constraint target words, while not differing from each other, were shorter than low- (.04) constraint target words (for converging evidence, see Lavigne et al., 2000, Experiment 2). High-constraint target words were also skipped more often than medium- and low-constraint target words. Rayner and Well concluded that "strong constraint is necessary to influence word skipping but not first-pass fixation times" (p. 507). Because other contrasts, for example, between low- and zero-constraint target words, have not been tested, it is still unclear how predictable a word needs to be in order to show predictability effects.

McDonald and Shillcock (2003a, 2003b) argued for a different kind of predictability effect, which they called *low-level predictability* or *forward transitional probability* (henceforth TP) effects. Forward transitional probability is the likelihood that word n will follow word $n - 1$. For example, the verb *accept* is followed more often by the word *defeat* than by the word *losses*. Therefore, the TP of *defeat* is higher than the TP of *losses*. McDonald and Shillcock (2003b) reported that this affects the fixation times on the noun, with longer fixation durations when the TP is lower. According to McDonald and Shillcock, TP effects reflect the availability of low-level statistical knowledge between adjacent words used in a bottom-up fashion. Of importance, McDonald and Shillcock argued that TP effects are independent from contextual predictability effects. Strength of TP, they suggested, exerts an immediate influence on the eye-movement record, with effects showing up in first-fixation, single-fixation, and gaze-duration measures. In contrast, contextual predictability effects reflect, according to McDonald and Shillcock, high-level, integrative, top-down processes, and show up mainly in "later" eye-movement measures. However, given that many studies (see Rayner, 1998, for a review) have demonstrated that predictability influences first-fixation durations and skipping rates (and the decision to skip a word must be made early in processing), it is difficult to maintain such a distinction. In addition, it is unclear why predictability, as assessed by a Cloze task, would reflect only higher order processes. While Cloze tasks provide an index of contextual constraint, it is possible that they are affected in turn by TP (i.e., TP affects the likelihood of completion). We return to this issue in the General Discussion section.

Evidence for TP effects comes from two studies: one involving an experiment in which TP was manipulated and contextual predictability was controlled (McDonald & Shillcock, 2003a), and one involving a corpus analysis of eye-movement data (McDonald & Shillcock, 2003b). McDonald and Shillcock (2003a) investigated verb–noun combinations such as *accept defeat/losses*. On the basis of the 100-million-word British National Corpus (Burnage & Dunlop, 1992), they calculated the probability of a specific noun following a verb (p[noun|verb] = frequency[verb, noun]/frequency[verb]). After selecting 48 pairs for which one member had a high TP (e.g., *defeat*) and the other a low TP (e.g., *losses*) for a specific verb (e.g., *accept*), a neutral preceding context was generated for each verb (e.g., *They simply will not*). Because the nouns were controlled for length and frequency, any differences in reading times can arguably be related to the difference in TP. McDonald and Shillcock (2003a) found a small TP effect for early reading times (significant for first-fixation duration, and nearly significant for single-fixation duration, with a trend in gaze durations), with 11–13 ms shorter fixation times for high-TP words compared with low-TP words, and no statistical differences in the amount of skipping. While this seems like clear evidence for the independence of TP effects, one concern with this study is that a Cloze task showed that the neutral contexts were not equally neutral for both types of items: Given the sentence fragments up to and including the verb, participants were 10 times more likely to complete them with the high-TP noun than the low-TP noun (.08 vs. .008). Hence, while the mean predictability scores were quite low, and the predictability contrast smaller than has been tested before, the fact that the predictability and the TP measures are confounded is potentially damaging to the claim that predictability and TP are independent. As indicated previously, low-level factors like TP might be part of regular predictability estimates. To maintain that TP and predictability are indeed independent, it is therefore crucial to show that TP effects can be found when predictability is held constant.

McDonald and Shillcock (2003b) examined eye-movement data while participants read excerpts from newspapers and calculated the TP of each word in the texts. Using multiple linear regression analyses, they showed independent effects of TP on both first-fixation and gaze durations, but not on the probability of skipping a word. These data are therefore in line with the experimental data of McDonald and Shillcock (2003a). They also found that the TP measure accounted equally well for content and function words.

However, it should be noted that the method did not allow them to control for a host of factors that could potentially have influenced reading times: word identity, syntactic role, and, indeed, the predictability of the word in context. A subset of this corpus also revealed first-fixation and gaze-duration effects of backward TP, which means that the likelihood of the word following the fixated word affects the early reading times on the fixated word, with shorter reading times when the next word follows the fixated word often. While forward and backward TP are usually highly correlated (McDonald & Shillcock, 2003b), they do not need to be the same. For example, *attach* might be followed often by *importance*, but that does not necessarily mean that *importance* is, relatively speaking, often preceded by *attach*. McDonald and Shillcock (2003b) take the backward TP effect as evidence of parafoveal-on-foveal effects (see Kennedy, 2000, for effects of parafoveal processing; though see Rayner, White, Kambe, Miller, & Liversedge, 2003, and Rayner, Warren, Juhasz, & Liversedge, 2004, for alternative accounts of such effects). Lavigne et al. (2000) also found shorter reading times on the fixation prior to the target word, though only when the target word was a high-frequency word and the launch site of the prior fixation was close to the target word. It should be noted that neither McDonald and Shillcock (2003b) nor Lavigne et al. (2000) tested the contextual predictability of the preceding word, which could have been the cause of the effects.

In short, McDonald and Shillcock (2003a, 2003b) provided some interesting data suggesting that the likelihood of two words occurring together in text may be exploited immediately by readers. Crucially, they maintain that TP effects are independent from predictability effects, and they provided two arguments for this claim (McDonald & Shillcock, 2003b): First, backward TP effects, which are dependent on the likelihood of the next word to occur, are difficult to reconcile with a measure of contextual integration; second, the verb–noun expressions preceded by a neutral context with only small differences in Cloze values still showed significant effects of TP (McDonald & Shillcock, 2003a).

These results have interesting implications, both at the theoretical level with respect to models of eye movements and at the more practical level of experimental design. At a theoretical level, they suggest that TP or low-level predictability at least partly drives the when component of eye movements (i.e., the decision of when to move the eyes to the next word, which plausibly reflects the time needed to identify the currently fixated word). Because some models of eye-movement control in reading, like the E-Z Reader model (Reichle, Rayner, & Pollatsek, 2003), rely heavily on frequency and predictability information as the prime inputs in making decisions regarding when to move the eyes, effects of TP, if real, should be incorporated into such models. One way to integrate forward TP, according to McDonald and Shillcock (2003b, pp. 1748–1749), is to replace the predictability component in the model's first stage (the "familiarity check") with a measure of TP. Incorporating backward TP, however, would be difficult in a sequential model like E-Z Reader. Thus, it is important to determine if TP effects indeed affect eye-fixation times. McDonald and Shillcock's results also suggest that TP effects are beyond a reader's conscious control as they are driven by mere statistical probabilities found in texts. These results are also important because they are, to our knowledge, the first potential evidence that at least one type of adjacent dependencies, namely, the co-occurrence frequency of specific verb–noun combinations, are reflected in reading times. Research on artificial grammar learning has shown that adults and children can learn adjacent dependencies between syllables, tones, and sequences of words in strings without formal instructions (e.g., Aslin, Saffran, & Newport, 1998; Gómez & Gerken, 1999; Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999; Saffran & Wilson, 2003), as well as some nonadjacent dependencies in certain restricted contexts (e.g., Gómez, 2002). McDonald and Shillcock's results would broaden the implications of these findings by showing that statistical dependencies are not restricted to the learning of artificial stimuli but are also brought into play during the processing of natural language.

At a practical level, these results could force researchers to take into account TP values when examining the eye-movement record in almost all cases in which a difference is expected between two conditions. For these reasons alone, an investigation of whether McDonald and Shillcock's (2003a) results replicate is important. However, we also wanted to extend their research and test whether TP effects are still discernable in more constraining contexts. If TP effects can be found even when predictability is high, then we can safely conclude that TP and predictability are independent and additive effects. If, however, TP effects disappear or are overridden when predictability increases, then we are faced with two possibilities: Either TP and predictability are not independent or TP effects are independent of predictability effects but only have measurable effects in particular types of context (see, e.g., Blutner & Sommer, 1988, on how sentence focus can override priming effects). To test the influence of predictability on TP, we placed verb–noun expressions that differed with respect to TP in neutral contexts and constraining contexts. We ran two experiments, the first using the same TP pairs as McDonald and Shillcock (2003a) and the second using a new set of items.

Experiment 1

The goal of Experiment 1 was twofold. Using the same TP item pairs as eye-movement measures as McDonald and Shillcock (2003a), we attempted to replicate their finding of TP affecting early eye-movement measures. By introducing constraining contexts in addition to neutral contexts, we wanted to examine the effects of predictability and determine whether predictability interacted with TP.

Method

Participants. Forty native American English-speaking students from the University of Massachusetts were paid $8 to participate in the experiment. All had normal or corrected-to-normal vision, and none participated in the pretest (see the *Materials* section).

Materials. We used the same 48 verb–noun pairs as McDonald and Shillcock (2003a; see Appendix A). The transitional probabilities were computed from the 100-million-word British National Corpus (Burnage & Dunlop, 1992). The mean number of occurrences of the verb–noun combinations in the corpus was 61.0 (range: 9–331) for the high-TP items and 1.6 (range: 0–7) for the low-TP items, $t(47) = 6.74$, $p < .001$. In terms of probability estimates, following McDonald and Shillcock's (2003a) definition (p[noun|verb] = frequency[verb, noun]/frequency[verb]), the mean TP values were .00989 and .00037 for high-TP nouns and low-TP nouns,

respectively.[1] We also calculated the backward TP values for the verbs (p[verb|noun] = frequency[verb, noun]/frequency[noun]): The mean backward TP for the verb when followed by a high-TP noun was .01089 and .00026 when followed by a low-TP noun. The nouns were controlled pairwise for frequency (log[10]-frequency = 3.0 for both high and low TP, based on the CELEX database; Baayen, Piepenbrock, & Van Rijn, 1993) and length (7.2 characters on average for both high and low TP). To respect American English spelling, we changed *colour* to *color* in one item. Analyses with this item excluded did not change the pattern of results.

The verb–noun pairs were preceded by either a contextually constraining context or a neutral context. The neutral contexts were similar to those from McDonald and Shillcock (2003a), though they were extended so that they had a comparable length to the constraining contexts. The Cloze test provided next indicated that the increased length of the neutral contexts did not affect predictability. A neutral spillover region, which was held constant within each item quartet, followed the critical noun. Examples can be found in (1a–d).

1a. As they cannot afford to lose the game, the team will not accept defeat even when they're far behind (constraining context, high TP: C–H).

1b. Since their mission is to make money, the manager will not accept losses even when it means pay cuts (constraining context, low TP: C–L).

1c. It is silly that they simply will not accept defeat even when they're far behind (neutral context, high TP: N–H).

1d. It is silly that they simply will not accept losses even when it means pay cuts (neutral context, low TP: N–L).

Contextual predictability was assessed by means of a modified Cloze test. A total of 41 participants completed sentence fragments consisting of the prior context up to and including the verb with the first word or words they could come up with. The mean Cloze values were .22 (range: .00–.73; SD = .20) for C–H, .06 (range: .00–.67; SD = .13) for C–L, .07 (range: .00–.70; SD = .14) for N–H, and .01 (range = .00–.23; SD = .03) for N–L. Note that the Cloze values for N–H and N–L were almost exactly the same as in McDonald and Shillcock's (2003a) Cloze test (.08 vs. .01), as were the ranges and standard deviations.

The experimental items were divided into four lists so that each participant saw 12 items in each condition and one version of each item. The 48 critical sentences were mixed with 85 filler sentences of various types. The sentences were presented in a fixed random order.

Procedure. Participants were tested individually. Eye movements were recorded using a Fourward Technologies Dual Purkinje Generation 5.5 eye tracker, which has an angular resolution of less than 10 min arc. All sentences were displayed on a single line or on two lines with a maximum length of 80 characters (the target regions all appeared on the first line). While viewing was binocular, only the right eye was monitored. Stimuli were displayed on a 15-in. NEC 4FG color monitor 61 cm from the participants' eyes. At this distance, 3.8 character positions equaled 1° of visual angle. Participants' gaze position was recorded every millisecond. A bite bar and forehead rest were used in order to minimize head movements.

Participants were encouraged to read the sentences carefully for understanding and to read at a normal rate. They pressed a button to remove a sentence when they had finished reading. Comprehension questions appeared after 50% of the items (balanced across conditions). Half of the questions required a yes response, half a no response. Accuracy was 92.5%.

At the beginning of the experiment, a calibration procedure was carried out. Calibration checks were performed before each sentence, and recalibration was done whenever necessary. The entire experiment lasted approximately 45 min, including a short break in the middle.

Analyses. An automatic software procedure pooled short contiguous fixations. Fixations that were less than 80 ms and within one character of another fixation were incorporated into one larger fixation. Fixations of less than 80 ms and not within one character space of another fixation were deleted as, presumably, readers hardly extract any information during these short fixations (see Rayner & Pollatsek, 1989). We also excluded fixations of more than 800 ms (e.g., Niswander, Pollatsek, & Rayner, 2000). In comparison, McDonald and Shillcock (2003a) did not report a lower cutoff and used a 700-ms upper cutoff.

We report analyses on three critical regions: the verb, the noun, and a spillover region that was defined as the next word if five characters long or more (including the space before the word). If the spillover word was shorter, the next two words were used. We included the spillover region because previous research (Calvo & Meseguer, 2002) has indicated that predictability effects are sometimes found immediately after reading the target word. While McDonald and Shillcock (2003a) reported only analyses on the noun, McDonald and Shillcock (2003b) argued that backward TP affects early reading times as well. In other words, the predictability of the fixated word given the following word should influence reading times in the same manner as forward TP influences reading times. Concretely, the argument is that *defeat* should be read faster than *losses* because *defeat* follows *accept* more often than *losses* follows *accept* (forward TP). In addition, *accept*, when followed by *defeat*, will be read faster than *accept* followed by *losses* (backward TP). Hence, according to this argument, backward TP effects should be discernable on the verb itself.

The same measures that McDonald and Shillcock (2003a) used are reported next: *first-fixation duration* (the duration of the first fixation on a word), *single-fixation duration* (the fixation time on a word when the word is fixated only once), *gaze duration* (the summed fixation times on a word during first-pass reading), and *probability of skipping*. The reported means and analyses of variance (ANOVAs) are based on reading times excluding 0-ms fixations (i.e., when a region was not fixated).

Results and Discussion

Prior to all analyses, sentences with major track losses, as a result of minor head movements or blinks, were excluded, as were sentences where two consecutive regions were skipped (less than 1% of the trials in total). For each measure and for each of the two regions, we subjected the data to separate Context (constraining context vs. neutral context) × TP (high TP vs. low TP) ANOVAs, treating participants (F_1) and items (F_2) as random effects. Table 1 presents the averages, using participant means, for first-fixation duration, single-fixation duration, gaze duration, and probability of skipping. We discuss the verb, noun, and spillover regions separately.

Verb region. If backward TP has an effect, we should find that the verbs with a high backward TP are read faster than verbs with a low backward TP. All the verbs in the C–H and N–H conditions had higher backward TP than the verbs in the C–L and N–L conditions (see the *Materials* section). No significant effects emerged on first-fixation duration analyses. The 11-ms difference between N–H and N–L on first-fixation duration did not yield

[1] These values are slightly different from the values quoted in McDonald and Shillcock (2003a), .01011 and .00038, which appears to be due to different versions of the British National Corpus. We feel that this is not of much significance because the discrepancy between the high- and the low-TP items is both substantial and comparable to McDonald and Shillcock (2003a). We thank Scott McDonald for his help in resolving this issue. Also note that McDonald and Shillcock (2003a), in addition to this simple TP measure, present a Bayesian model to analyze TP effects.

Table 1
Experiment 1: Mean Reading Times and Probability of Skipping

	Verb region		Noun region		Spillover region	
Measure	*M*	*SE*	*M*	*SE*	*M*	*SE*
First-fixation duration						
C–H	279	7.4	283	6.5	279	9.7
C–L	273	6.2	292	7.2	270	7.0
N–H	278	7.4	294	6.8	281	6.8
N–L	289	7.8	300	6.4	289	6.4
Single-fixation duration						
C–H	283	7.6	290	6.9	288	11.1
C–L	285	6.6	299	7.4	274	7.3
N–H	284	7.7	301	7.3	282	7.2
N–L	297	8.5	309	7.3	294	6.8
Gaze duration						
C–H	297	6.1	315	6.6	313	10.1
C–L	292	5.2	327	7.4	301	7.9
N–H	304	6.6	318	7.1	317	10.0
N–L	313	8.1	338	6.1	327	8.3
Probability of skipping						
C–H	.17	.02	.09	.02	.38	.03
C–L	.19	.02	.08	.01	.29	.03
N–H	.16	.02	.06	.01	.29	.03
N–L	.16	.02	.09	.02	.28	.03

Note. Reading times are in milliseconds. C–H = constraining context/high-transitional probability; C–L = constraining context/low-transitional probability; N–H = neutral context/high transitional probability; N–L = neutral context/low-transitional probability.

significance in a planned comparison (ps $> .10$). Likewise, no significant effects emerged for the single-fixation duration. The 13-ms difference between N–H and N–L did not yield significance in a planned comparison (ps $> .10$). However, the gaze-duration data yielded a significant effect of context: When the verbs appeared in a constraining context, they were read faster than when they followed a neutral context: $F_1(1, 39) = 8.23$, $p < .01$; $F_2(1, 47) = 5.43$, $p < .03$. The 9-ms difference between N–H and N–L was not significant (ps $> .20$). No significant differences were found for the probability of skipping.

In short, there was no sign of a backward TP effect. The context effect that was found in the gaze-duration data fits into the idea that more predictable words are read faster than less predictable ones, though it should be noted that the Cloze values reported previously are the predictability ratings for the noun rather than the verb. We therefore carried out a Cloze test with 24 new participants for the sentence fragments excluding the verb. Each participant saw one constraining and one neutral fragment for each verb so that each context was completed by 12 participants. The results indicated that predictability for the constraining contexts was four times larger than the predictability for the neutral contexts (constraining: .04; neutral: .01), which might have been the source of the context effect.

Noun region. First-fixation duration analyses showed that the nouns were read faster in a constraining context than in a neutral context: $F_1(1, 39) = 5.70$, $p < .03$; $F_2(1, 47) = 5.07$, $p < .03$. High-TP nouns also tended to be read faster than low-TP nouns, though this was only significant in the participants analysis: $F_1(1, 39) = 4.37$, $p < .05$; $F_2(1, 47) = 2.46$, $p = .12$. The interaction was not significant (both Fs < 1). Single-fixation durations showed the same pattern: a significant effect of context, $F_1(1, 39) = 5.93$, $p = .02$; $F_2(1, 47) = 5.07$, $p < .03$, an effect of TP significant by participants only, $F_1(1, 39) = 6.06$, $p < .02$; $F_2(1, 47) = 1.36$, $p = .25$, and no interaction (Fs < 1). Finally, gaze-duration analyses yielded a significant effect of TP: $F_1(1, 39) = 15.32$, $p < .001$; $F_2(1, 47) = 6.99$, $p = .01$, but no other effects. The analyses on the probability of skipping showed no significant effects.

Spillover region. First-fixation durations were longer in the neutral context conditions than in the constraining context condition, though this was fully significant only in the items analysis: $F_1(1, 39) = 3.40$, $p < .08$; $F_2(1, 46) = 5.16$, $p < .03$.[2] The single-fixation duration analyses revealed no significant effects. Gaze-duration analyses showed a significant effect of context, with constraining contexts leading to shorter spillover reading times: $F_1(1, 39) = 4.45$, $p < .05$; $F_2(1, 46) = 7.51$, $p < .01$. The probability of skipping analyses showed a significant effect of context, $F_1(1, 39) = 9.45$, $p < .01$; $F_2(1, 47) = 7.02$, $p = .01$, a marginal effect of TP, $F_1(1, 39) = 3.57$, $p < .07$; $F_2(1, 47) = 7.09$, $p = .01$, and an interaction, $F_1(1, 39) = 4.14$, $p < .05$; $F_2(1, 47) = 4.54$, $p < .04$. Planned comparisons showed that the spillover region in the C–H condition was skipped more often than in the other three conditions (all ps $< .03$), which did not differ from each other (all ts < 1).

In summary, our data are generally very much in line with McDonald and Shillcock's (2003a) data: Nouns that follow a particular verb more often (high-TP nouns) were read faster than control nouns that follow the same verb less often (low-TP nouns). While McDonald and Shillcock (2003a) observed this pattern most

[2] There was one missing cell in the items analysis.

strongly in their first fixation data, we observed it most strongly in the gaze-duration data. Because the trends in our other measures were strong, one should be cautious about drawing too strong of a conclusion from this difference in time course, though it is worth pointing out that if TP effects truly reflect automatic low-level bottom-up processes, we would probably have expected the strongest effects to turn up on the earliest measures (first- and single-fixation durations; see McDonald & Shillcock, 2003a, 2003b). In addition to TP effects, we also found significant early effects of context, with constraining contexts leading to shorter reading times than neutral contexts on both the verb and the noun regions. This result suggests that our contextual manipulation was indeed strong enough and that predictability effects can be observed with quite small mean Cloze differences at the lower end of the scale. They also corroborate previous findings (Rayner & Well, 1996) of predictability effects that were found for early reading time measures and show that these effects are not restricted to late processing measures like rereading time and number of regressions back to a word (*contra* Calvo & Meseguer, 2002). The lack of an interaction between context and TP indicates that our context manipulation did not override the TP effect and that TP effects could operate independently from predictability effects. The skipping data are also in line with McDonald and Shillcock (2003a) in the sense that no differences in skipping probability were found on the verb and the noun. However, the interaction observed in the spillover region indicates that when the context is constraining and the TP of the noun is high, skipping of the next region is more likely. This interaction is unexpected if TP and predictability are acting independently from each other, but can be explained if one assumes that they can reinforce each other.

As indicated previously, the Cloze values for the low-TP nouns preceded by a neutral context were substantially lower than the Cloze values for the high-TP nouns in a neutral context. Because our data suggest that even relatively small differences in mean predictability ratings can lead to context effects, this difference in the neutral context conditions could have influenced the fixation times. In fact, rather than comparing N–H to N–L, which have predictability ratings of .07 and .01, one could argue that a better comparison would be C–L to N–H, in which the predictability ratings, standard deviations, and ranges are much closer. When comparing these two conditions, no significant differences emerged: for first- and single-fixation durations, ts < 1; for gaze duration, $t_1(39) = 1.52$, $p > .13$; $t_2(47) = 1.13$, $p > .26$. Hence, this suggests that when mean predictability is controlled, TP effects might no longer be observable.

However, mean predictability is potentially misleading as it does not indicate item variability. If TP effects are independent from predictability effects, we should expect a comparable speeding up for high-TP items, whether the discrepancy in Cloze values is small or large. However, if TP effects are akin, or part of, general predictability effects, we would expect that the difference in reading times will be largest for item pairs that differ most in predictability. A similar argument can be made for item pairs in the constraining context conditions: shorter reading times when the difference in predictability is large. A straightforward test of this possibility can be obtained by correlating the difference in reading times with the difference in predictability, which showed a near-significant effect in the predicted direction for the gaze durations for the neutral contexts ($r = -.22$, $p = .07$, one-tailed), though not for the constraining contexts ($r = -.08$, $p = .30$). However, Hyönä's (1993) and Rayner and Well's (1996) findings discussed previously indicate that predictability effects are very likely not linear, especially not for more constraining contexts, and correlations might not be the right type of analysis for that reason. We therefore decided to split the data into two subsets: a subset of matched items (i.e., items with comparable Cloze values) and a subset of mismatched items. We did this for the neutral and constraining context conditions separately. Note that, because of the pairwise controlling of the target nouns, frequency and length are still well controlled in these subset analyses. For the neutral context conditions, we identified 26 matched item pairs for which the N–H and N–L conditions all had a Cloze value of .00 and 21 mismatched item pairs for which the Cloze value of N–H was higher than the value of N–L (.12 and .00, respectively). One item for which the Cloze value was higher for N–L than for N–H was deleted from the analyses. For the matched items, the gaze durations on the noun for N–H were 15 ms less than for N–L, a nonsignificant difference, $t_1(39) = 1.36$, $p > .18$; $t_2(25) = 1.36$, $p > .18$; for the mismatched items, this difference was 24 ms, $t_1(39) = 2.55$, $p < .02$; $t_2(20) = 3.16$, $p < .01$.

The items in the constraining context conditions, after taking out 2 items for which the Cloze of the C–L condition was at least .20 higher than the C–H condition, were arranged in two subsets based on a median split. The matched subset consisted of 29 item pairs for which the difference in Cloze values was equal or less than the .18 median (Cloze C–H = .11, Cloze C–L = .04). The mismatched subset contained 17 items for which the difference in Cloze values was higher than the median (Cloze C–H = .42, Cloze C–L = .01). For the matched items, the 3-ms advantage for the C–H condition in gaze durations was not significant (both ts < 1). For the mismatched items, the 31-ms advantage for C–H compared with C–L was significant in the participants analysis, $t_1(39) = 2.34$, $p < .03$; $t_2(16) = 1.55$, $p = .14$. Together, the results of the subset analyses indicate that when the predictability of the low- and high-TP items is controlled, no significant TP effects could be found. It is therefore quite possible that apparent TP effects are actually due to Cloze differences.

McDonald and Shillcock (2003a, 2003b) also argued that TP effects are more likely to occur when the fixation preceding the one on the noun happened to be close to the noun; that is, when the launch site for the target fixation was close. They argued that this would induce an effect of TP because parafoveal preview is greater at short launch distances and because converging evidence from parafoveal and statistical (i.e., transitional probability) information will facilitate processing. To test this possibility, we divided the gaze-duration data in two groups: close launch sites (i.e., launch sites of 1 to 4 characters to the left of the space before the noun) and far launch sites (5–15 characters to the left). We then added launch distance as an extra factor in the ANOVAs and tested for an interaction between launch distance (close vs. far) and TP. Launch distance did not interact with TP (both Fs < 1). For the first-fixation data, this interaction was marginally significant in the participants analysis, $F_1(1, 37) = 3.59$, $p = .07$, but not in the items analysis, $F_2(1, 44) = 1.10$, $p = .30$. However, when analyzing the two conditions best matched for predictability (C–L and N–H), the interaction was far from significant (both Fs < 1).

In conclusion, while our data at first sight replicate McDonald and Shillcock's (2003a) finding that TP effects can be found in

very early reading time measures, there are reasons to believe that these effects might actually be due to predictability. First, no significant differences were found when comparing two conditions that had comparable Cloze values (C–L and N–H). Second, when item pairs were split into subsets that were either matched for predictability or not matched, significant TP effects were found only for unmatched pairs, both for the neutral and constraining context conditions. Hence, it might be the case that McDonald and Shillcock's (2003a) results reflected differences in predictability rather than transitional probability, or that TP is one factor influencing predictability judgments. Experiment 2 was therefore set up with two intentions in mind: We wanted to determine whether TP effects are still discernable when predictability for the neutral and the constraining contexts is better controlled, and whether predictability can be observed in early reading measures for items with relatively small Cloze differences between conditions.

Experiment 2

The results of Experiment 1 raise the possibility that apparent TP effects might have been caused by differences in predictability, and that a constraining context can speed up reading rates even when the mean difference in predictability is fairly small. Experiment 2 used the same design as Experiment 1, but we constructed new items so that the Cloze values of the item pairs were much better matched. This way, if effects of TP are found, in addition to predictability effects, then it will be difficult to argue that they are caused by mismatches in predictability. On the other hand, if no effects or only predictability effects are observed, then this result will further support the notion that TP effects are not independent from predictability effects.

Another motivation for Experiment 2 was our concern with a number of items from Experiment 1 (and, thus, McDonald & Shillcock's, 2003a, experiment) that might have exhibited TP effects for reasons unrelated to a difference in TP or predictability. For example, a number of item pairs involved verbs such as *cast* that were semantically ambiguous, as in *cast girls* and *cast doubt*. As shown by Pickering and Frisson (2001), homonymous verbs are resolved differently from other verbs (see also Binder & Rayner, 1998; Duffy, Morris, & Rayner, 1988; Rayner, Pacht, & Duffy, 1994; Sereno, Pacht, & Rayner, 1992, on the processing of homonymous nouns). Hence, because the "typecast" interpretation of *cast* is the less frequent meaning of *cast*, the low-TP noun *girls* might require more processing than a noun compatible with the more frequent interpretation of *cast*, not because its TP is lower but rather because it is paired with the verb's subordinate interpretation. Other items might have been problematic for a different reason. Expressions such as *pay tribute* and *save face*, which are high-TP items, are also idiomatic. Previous research on idioms (e.g., Gibbs & Gonzales, 1985; Swinney & Cutler, 1979) suggests that idiomatic interpretations are processed faster than literal ones, possibly because the former can be stored as a unit while the latter needs to be assembled. We therefore excluded clearly idiomatic expressions as much as possible in the current experiment.

Method

Participants. Thirty-six American English-speaking students from the University of Massachusetts participated in the experiment for either credit or money.[3] All had normal or corrected-to-normal vision, and none participated in the pretest (see the *Materials* section).

Materials. We tested a total of 56 item quartets, 30 of which used the same verb as in Experiment 1 (see Appendix B). Some of these verbs were followed by the same noun as in Experiment 1, while others were paired with different nouns, and contexts were rewritten when appropriate. The same restrictions on length, frequency, and TP were used as in Experiment 1. The mean number of occurrences, as assessed by the British National Corpus, was 51.0 (range: 12–331) for the high-TP items and 1.2 (range: 0–7) for the low-TP items, $t(55) = 6.71$, $p < .001$. Mean TP estimates were .00677 for high-TP nouns and .00018 for low-TP nouns. Mean backward TP estimates for the verb were .01204 when followed by a high-TP noun and .00054 when followed by a low-TP noun. Frequency and length of the nouns were controlled (log[10]-frequency = 2.9 for high-TP items and 2.8 for low-TP items; mean length = 7.2 characters for both high- and low-TP items).

The design was the same as in Experiment 1, and contextual predictability was assessed by a modified Cloze task. A total of 56 participants completed sentence fragments consisting of the prior context up to and including the verb with the first word or words they could come up with. The mean predictability values were .20 for the C–H items (range: .0–.85; $SD = .23$), .18 for C–L (range: .0–.79; $SD = .20$), .02 (range: .0–07; $SD = .03$) for N–H, and .01 (range: .0–.07; $SD = .02$) for N–L. Hence, the imbalance in predictability ratings between the high- and the low-TP items that was present in Experiment 1 was removed in Experiment 2. The experimental items were divided into four lists so that each participant saw 14 items in each condition and one version of each item. The order of presentation was randomized for each participant. In addition to the 56 experimental items, each list contained 84 filler items of various types. As the filler items consisted of two sentences, we added a second sentence to the critical items. Because the critical sentence was always the first sentence and because we are interested in the reading times before the participants reached the second sentence, this could not have affected our results.

Procedure and analyses. Procedure, apparatus, and analyses were the same as in Experiment 1. The spillover region was defined the same as in Experiment 1. The entire experiment lasted about 50 min, including a short break in the middle. The accuracy on the comprehension questions was 91.3%.

Results and Discussion

Prior to all analyses, sentences with track losses were excluded (2.0% of the trials). Table 2 presents the averages, using participant means, for first-fixation duration, single-fixation duration, gaze duration, and probability of skipping. We discuss the three regions—verb, noun, and spillover—separately.

Verb region. There were no significant or near-significant effects in the first-fixation duration, single-fixation duration, gaze-duration, or probability-of-skipping analyses (all ps $> .14$).

Noun region. There were no significant or near-significant effects in the first-fixation duration data: The 7-ms difference between C–H and C–L was not significant, $t_1(35) = 1.51$, $p > .14$; $t_2(55) < 1$. No significant or near-significant effects were observed in the single-fixation duration data, and the 10-ms difference between C–H and C–L was not significant, $t_1(35) = 1.58$, $p > .12$; $t_2(55) < 1$. The gaze-duration data showed an effect of context, with constraining contexts leading to shorter reading times, $F_1(1, 35) = 3.82$, $p < .06$; $F_2(1, 55) = 4.85$, $p < .04$, but no effect of

[3] Four additional participants (one from each counterbalancing list) were deleted because of their high skipping rates.

Table 2
Experiment 2: Mean Reading Times and Probability of Skipping

	Verb region		Noun region		Spillover region	
Measure	*M*	*SE*	*M*	*SE*	*M*	*SE*
First-fixation duration						
C–H	288	6.8	300	6.9	297	7.5
C–L	283	6.6	307	7.6	303	8.8
N–H	292	7.8	307	6.8	313	8.4
N–L	288	8.3	309	7.0	313	9.1
Single-fixation duration						
C–H	292	7.6	305	7.2	304	8.6
C–L	289	7.5	315	7.8	314	10.6
N–H	300	8.8	313	6.8	321	10.1
N–L	295	8.7	315	7.4	322	11.8
Gaze duration						
C–H	317	8.7	328	7.3	340	8.8
C–L	315	8.8	334	7.4	347	10.5
N–H	325	9.8	342	8.1	354	10.6
N–L	317	8.7	339	8.0	359	12.9
Probability of skipping						
C–H	.12	.02	.07	.01	.17	.02
C–L	.11	.02	.07	.01	.18	.03
N–H	.13	.02	.06	.01	.17	.03
N–L	.11	.02	.06	.02	.15	.02

Note. Reading times are in milliseconds. C–H = constraining context/high-transitional probability; C–L = constraining context/low-transitional probability; N–H = neutral context/high-transitional probability; N–L = neutral context/low-transitional probability.

TP or an interaction (all Fs < 1). The 6-ms difference between C–H and C–L was not significant (both ts < 1). The probability-of-skipping analyses showed no significant or near-significant effects.

Spillover region. First-fixation duration analyses revealed a significant effect of context, with shorter reading times on the spillover region when preceded by a constraining context: $F_1(1, 35) = 4.60$, $p < .04$; $F_2(1, 55) = 4.27$, $p < .05$. Trends in the same direction were found for single-fixation duration, $F_1(1, 35) = 3.04$, $p = .09$; $F_2(1, 55) < 1$, and gaze duration, $F_1(1, 35) = 3.20$, $p < .09$; $F_2(1, 55) = 3.26$, $p < .08$, analyses. There were no significant effects of TP and no significant interactions (all ps $> .29$). There were no effects in the skipping analyses.

The results of Experiment 2 are easy to summarize: An effect of context (i.e., a predictability effect) was found for the gaze duration on the noun and the first-fixation duration on the spillover region, but no evidence was found that TP exerted an influence on early reading times. In fact, the reading times on N–H and N–L, the conditions akin to the ones used by McDonald and Shillcock (2003a), are remarkably alike, with differences of no more than a 5-ms difference on any measure. Given that the frequency of the high-TP nouns was slightly higher than the low–TP nouns and that the Cloze values for N–H were slightly higher than the Cloze values for N–L, this difference is negligible. The main conclusion to be drawn from these results is that, with predictability more tightly controlled, differences in TP did not affect early reading times for our new set of items. In contrast, we did identify a context effect. While the context effect was observed in the first- and single-fixation duration data in Experiment 1, this effect was restricted to the gaze durations in the present experiment. We are not sure why this is the case, nor do we know whether this reflects any significant processing differences because all three measures are generally considered early measures of processing.

As in Experiment 1, we split the data into matched and mismatched subsets. For the neutral context conditions, we identified 20 items for which the Cloze value was .00 (matched items) and 20 items for which the N–H Cloze value was higher than the N–L Cloze value (.05 and .00, respectively). Note that this contrast is much less than in Experiment 1 as a result of better controlling of predictability. The only significant effect on the noun, and only for the participants analysis, was found for the first-fixation data for the mismatched subset, with a 23-ms advantage for the N–H condition, $t_1(35) = 2.72$, $p = .01$; $t_2(19) = 1.66$, $p = .11$. For the constraining context conditions, we followed the same procedure as in Experiment 1. After taking out item pairs for which the Cloze of the C–L condition was at least .20 higher than the C–H condition (12 items), a matched and mismatched subset was distinguished on the basis of a median split. The matched subset consisted of 22 items with minor differences in Cloze values (.07 and .10 for C–H and C–L, respectively). No significant effects were observed for this subset (all ts < 1). The mismatched subset consisted of 22 items for which the difference in Cloze value was higher than the median (.38 and .13 for C–H and C–L, respectively). Again, this contrast is much less than in Experiment 1. No significant effects were found for the planned comparisons (all ps $> .15$). We also examined whether launch distance (close: 1 to 4 characters to the left of the space before the noun vs. far: 5–15 characters) and TP interacted for the gaze-duration data. As in Experiment 1, no significant interaction was observed: $F_1(1, 35) = 3.03$, $p = .09$; $F_2(1, 55) = 2.33$, $p > .13$. The same analysis for the first-fixation duration data revealed a significant interaction for the participants analysis, $F_1(1, 35) = 4.08$, $p = .05$, but not for the

items analysis, $F_1(1, 35) = 2.13$, $p = .15$. While this seems to suggest that TP effects might be slightly stronger when the launch site is close to the target word, one has to be very careful with this interpretation, as these analyses are based on a much reduced number of observations per cell, and the built-in experimental controls (length, frequency, predictability) are at least partly lost in this procedure (which was not the case in the subset analyses). Consequently, some items might have disproportionately influenced mean reading times for factors not related to TP. For example, taking out the 3 items for which the predictability of the C–H condition was at least .50 more than the predictability of the C–L condition made the interaction disappear: $F_1(1, 35) = 2.35$, $p = .13$; $F_2(1, 55) = 1.62$, $p = .24$. In short, we found no convincing evidence that TP effects depend on launch distance, though further research explicitly designed to test this possibility is warranted.

General Discussion

Experiment 1 showed effects of TP and predictability in the early reading time measures. The lack of an interaction suggests that TP has an effect independent of predictability. However, comparisons of conditions that were better matched on strength of predictability and analyses of subsets cast doubt on this interpretation, and suggest that the purported TP effects could be traced back to regular predictability effects. Experiment 2 followed up on this suggestion and showed that, when predictability is well controlled, TP effects disappeared while regular predictability effects were still observed. Neither of the experiments showed significant effects of backward TP (i.e., that the probability of the upcoming word affects reading times on the fixated word). This result is a clear lack of a parafoveal-on-foveal effect (see Rayner et al., 2003, 2004).

As we discussed in the introduction, McDonald and Shillcock (2003b) provided two arguments for why TP effects are separate from predictability effects. Their first argument concerned their finding of backward TP effects, which are supposed to be incompatible with the idea of predictability being a measure of high-level contextual integration. In both experiments, we failed to replicate this finding, even though the differences in backward TPs were substantial. There might be several reasons for this inconsistency. McDonald and Shillcock (2003b) obtained reading times from a corpus they compiled by having participants read excerpts from newspaper articles and then calculated forward and backward TPs for each word in the corpus. They then selected a set of words with high backward TP and a set with low backward TP, and showed shorter reading times for the first set compared with the second set. Hence, it might be possible that the difference in backward TP for the high and low items was greater than in our experiments. However, by examining backward probability this way, they were unable to control other variables such as word type, syntactic category, and, indeed, predictability.

Their second argument in favor of the independence of TP effects was their finding (McDonald & Shillcock, 2003a) that TP effects occurred in neutral sentences with low mean Cloze values. As discussed previously, the additional analyses of Experiment 1 and the results of Experiment 2 strongly suggest that this effect was not independent from predictability effects. This outcome also affected our original plan to test whether constraining contexts could override TP effects. If one believes that TP effects are a reflection of low-level statistical probabilities independent from high-level cognitive predictability or expectancy, constraining contexts should not have an influence on TP. However, because the TP effects for the neutral context conditions in Experiment 1 were questionable, and nonexistent in Experiment 2, the question whether TP effects can be overridden by context ceases to be a concern.

So, why is it that McDonald and Shillcock (2003a) found effects of TP while we failed to replicate them with a larger set of items that were strictly controlled for predictability (Experiment 2)? There are a number of factors that could possibly have played a role, either alone or in combination. The main difference between the two experiments was the predictability scores for the neutral context conditions. While the mean Cloze values for the low-TP items were comparable in both experiments (both .01), the mean Cloze values for the high-TP items were dissimilar: .08 in McDonald and Shillcock's (2003a) experiment versus .02 in Experiment 2. Another difference was that we excluded a number of fixed expressions. If these expressions behaved like idioms, then they might have been processed faster than the nonidiomatic member of the item pair (e.g., Gibbs & Gonzales, 1985; Swinney & Cutler, 1979). We discuss two, not mutually exclusive, ways of how higher Cloze values could have influenced their results.

First, it is possible that relatively small differences in Cloze can lead to predictability effects, especially at the low end of the scale. While relatively large differences at the medium-high end of the scale do not seem to have much impact on fixation times (Hyönä, 1993; Rayner & Well, 1996), no one has systematically investigated the possibility that predictability effects can be observed between words that have low Cloze and words with zero Cloze. It might be possible therefore that predictability effects are more like frequency effects, for which it has been found that small differences in frequencies at the low end of the scale have a bigger effect than at the high end of the scale (see also Gernsbacher, 1984). There is some evidence supporting this view. The results of Rayner and Well (1996) indicate that predictability is not a linear function, as no effect was found between medium- and high-constraining words on fixation time measures while the predictability difference was as high as between medium- and low-constraining words. In addition, the subset analyses of Experiments 1 and 2 showed that predictability effects were more reliable for the mismatched neutral context conditions than for the mismatched constraining context conditions, even as the predictability differences were larger for the latter subsets. Obviously, the hypothesis of predictability effects at the low end of the scale needs more experimental support, though we see no reason at this moment to reject it out of hand.

The second explanation for why the differences in Cloze values between McDonald and Shillcock (2003a) and Experiment 2 might have led to different outcomes is related to this view. Rayner and Well (1996) compared high- and medium-predictable words with low-predictable ones and found shorter gaze durations for both high- and medium-predictable words. A closer look at the predictability values for McDonald and Shillcock's (2003a) items revealed that a number of them had medium- to high-predictability values, even in what was considered a neutral context. When only those items were taken into account that had Cloze values of zero, no effects became apparent in our experiments. These results,

together with the almost identical early reading times for the neutral context conditions of Experiment 2, strongly suggest that McDonald and Shillcock's (2003a) results might have reflected differences in strength of predictability rather than differences in TP.

We do not want to argue that transitional probabilities have no significance whatsoever in reading, though we question whether TP effects are truly independent of what has traditionally been considered predictability effects. The finding that there is a slight bias, as can be deduced from the Cloze values, to complete a semantically neutral context, such as "It is silly that they simply will not accept," with a high-TP noun such as *defeat* might indicate that some kind of record is being kept of how often *accept* is followed by *defeat* (though it is equally possible that this bias is unrelated to TP and merely reflects the predictability of the sentence fragment as a whole). If this is the case, then contingency statistics might be one of several sources participants in a modified Cloze task can rely on in order to complete a sentence fragment. Presumably such statistics are also relevant to people's choices of words in normal production. In this sense, predictability measures, as estimated by means of a modified Cloze task, capture both low-level and high(er)-level probabilities. Because other factors like, for example, the degree of fit of a word in a larger context influence predictability ratings as well, it should be possible to find predictability effects without TP effects, though the reverse would be difficult or impossible. This is also exactly what our data show.

This article has addressed the role of TP and predictability on eye movements in reading. We know that eye movements provide an indication of language processing (Rayner, 1998) because they are affected by many cognitive variables (e.g., word frequency). It would therefore be interesting to see whether TP and predictability affect other measures of language processing. For instance, studies have investigated the effects of local and global context on processing difficulty (e.g., Hess, Foss, & Carroll, 1995; Schwanenflugel & White, 1991; Van Berkum, Hagoort, & Brown, 1999), and similar experiments could investigate whether TP has an independent effect on reaction time measurements of spoken or written language comprehension.

To conclude, our results question whether TP effects are truly independent of what is generally considered predictability effects. They are, however, compatible with an account that considers transitional probabilities as one component among several that influences predictability. This way, TP effects provide us with a better understanding of what can drive predictability ratings, though they do not provide evidence for a separate processing stage.

References

Altarriba, J., Kroll, J., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24,* 477–492.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9,* 321–324.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database* [CD]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364–390.

Binder, K., Pollatsek, A., & Rayner, K. (1999). Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1162–1172.

Binder, K., & Rayner, K. (1998). Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review, 5,* 271–276.

Blutner, R., & Sommer, R. (1988). Sentence processing and lexical access: The influence of the focus-identifying task. *Journal of Memory and Language, 27,* 359–367.

Burnage, G., & Dunlop, D. (1992). Encoding the British National Corpus. In J. M. Aarts, P. de Haan, & N. Oostdijk (Eds.), *English language corpora: Design, analysis, exploitation* (pp. 79–95). Amsterdam: Rodopi.

Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contributions of visual, lexical, and contextual factors. *The Spanish Journal of Psychology, 5,* 66–77.

Cole, R. A., & Perfetti, C. A. (1980). Listening for mispronunciations in a children's story: The use of contexts by children and adults. *Journal of Verbal Learning and Verbal Behavior, 19,* 297–315.

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology, 16,* 79–103.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical decision and fixation times in reading. *Journal of Memory and Language, 27,* 429–446.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20,* 641–655.

Fischler, I., & Bloom, P. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior, 18,* 1–20.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113,* 256–281.

Gibbs, R., & Gonzales, G. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition, 20,* 243–259.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431–436.

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70,* 109–135.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General, 124,* 62–82.

Hyönä, J. (1993). Effects of thematic and lexical priming on readers' eye moments. *Scandinavian Journal of Psychology, 34,* 293–304.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior, 23,* 612–624.

Kennedy, A. (2000). Parafoveal processing in word recognition. *The Quarterly Journal of Experimental Psychology, 53A,* 429–455.

Kutas, M., Lindamood, T. E., & Hilliard, S. A. (1984). Word expectancy and event-related potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 217–234). Hillsdale, NJ: Erlbaum.

Lavigne, F., Vitu, F., & d'Ydewalle, G. (2000). The influence of semantic context on initial landing sites in words. *Acta Psychologica, 104,* 191–214.

McClelland, J. M., & O'Regan, J. K. (1981). Expectations increase the benefit derived from parafoveal visual information in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 634–644.

McDonald, S. A., & Shillcock, R. C. (2003a). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science, 14,* 648–652.

McDonald, S. A., & Shillcock, R. C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research, 43,* 1735–1751.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes, 15,* 389–420.

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 556–573.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 720–732.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language, 33,* 527–544.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Hillsdale, NJ: Erlbaum.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). Eye movements and plausibility effects in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1290–1301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3,* 504–509.

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: North-Holland.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–526.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70,* 27–52.

Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy, 4,* 273–284.

Schwanenflugel, P. J., & LaCount, K. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 344–354.

Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language, 24,* 232–252.

Schwanenflugel, P. J., & White, C. R. (1991). The influence of paragraph information on the processing of upcoming words. *Reading Research Quarterly, 26,* 160–177.

Sereno, S. C., Pacht, J. M., & Rayner, K. (1992). The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science, 3,* 296–300.

Stanovich, K. E., & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition, 7,* 77–85.

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 658–672.

Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18,* 523–534.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11,* 657–671.

Appendix A

Experiment 1: Items

For each item, the constraining context/high-transitional probability (C–H) version appears first (a), followed by the constraining context/low-transitional probability (C–L) version (b). The (c/d) version is the neutral context, which was the same for the high-transitional probability (N–H) and the low-transitional probability (N–L) versions. The rest of the sentence was the same as in the constraining context versions, starting from the noun. The / symbols delimit the regions of analysis, the // symbols indicate a line break The space before the first word in a region of analysis was included in the analysis.

(a) As they cannot afford to lose the game, the team won't/ accept/ defeat/ even/ when// they're far behind.
(b) Since their mission is to make money, the manager won't/ accept/ losses/ even/ when// it means pay cuts.
(c/d) It is silly that they simply won't/ accept/ . . .

(a) To bring new employees up to speed, the firm will/ provide/ training/ on a/ real// system.
(b) After training in theoretical issues, the firm will provide/ practice/ on a/ real// system.
(c/d) We are very delighted to hear that the firm will/ provide/ . . .

(a) If you look through the binoculars, you might/ catch/ sight/ of that/ famous pack of// lions.
(b) If they switch on their TV now, they might/ catch/ parts/ of that/ famous series on// lions.
(c/d) If you're a little bit lucky, you might/ catch/ . . .

(a) Because environmental laws get weakened, activists/ express/ concern/ about/ what// the best course of action is.
(b) By means of writing letters to the editor, activists/ express/ opinion/ about/ what// they feel strongly about.
(c/d) It is not that surprising that activists/ express/ . . .

(a) Since John is the supervisor, he is supposed to/ monitor/ progress/ so that/ the// financiers will not get upset.
(b) Since John has huge debts, he is supposed to/ monitor/ spending/ so that/ the// financiers will not get upset.
(c/d) The man with the clipboard is supposed to/ monitor/ . . .

(a) By signing this official waiver, we/ grant/ permission/ to anyone/ who is in need// of it.
(b) By the authority of a police force, we/ grant/ protection/ to anyone/ who is in need// of it.
(c/d) Following the executive decision, we/ grant/ . . .

(a) Wanting to curb excesses, politicians plan to/ introduce/ legislation/ in a/ very// forward manner.
(b) Wanting the jury to see the horror, they plan to/ introduce/ photographs/ in a/ very// forward manner.
(c/d) It was rumored that the gentlemen plan to/ introduce/ . . .

(a) Surrounded by DEA agents, Jeff could not/ resist/ arrest/ and quickly/ gave up.
(b) As Jeff loves dairy products, he could not/ resist/ cheese/ and quickly/ gave in.
(c/d) As his partner had predicted, Jeff could not/ resist/ . . .

(a) After his stupid remark, Mike tried to/ save/ face/ in a/ last ditch attempt.
(b) After losing most of his money, Mike tried to/ save/ part/ in a/ last ditch attempt.
(c/d) Of those remaining, Mr. Robinson tried to/ save/ . . .

(a) By its popular policies, the party was able to/ attract/ support/ from/ a diverse// range of groups.
(b) By phoning people at home, the party was able to/ attract/ members/ from/ a diverse// range of groups.
(c/d) Quite unexpectedly, only one party was able to/ attract/ . . .

(a) The X-rays taken from the man's chest indeed/ show/ signs/ of a/ very unusual// affliction.
(b) The pictures of the mansion's interior indeed/ show/ doors/ of a/ very unusual// style.
(c/d) To our surprise, several of the photos indeed/ show/ . . .

(a) Because we now have all the facts, we must/ draw/ conclusions/ from/ the private// report.
(b) Because we need new surgeons, we must/ draw/ specialists/ from/ the private// hospital.
(c/d) In fact, it might turn out that we must/ draw/ . . .

(a) With the bartenders on strike, they will not/ serve/ drinks/ until/ the end of// next week.
(b) As the station is being renovated, it will not/ serve/ trains/ until/ the end of// next week.
(c/d) The notice says that this station will not/ serve/ . . .

(a) As Mr. Robertson is the key witness, he will/ give/ evidence/ in front/ of a// packed audience.
(b) As Rob is the best football coach around, he will/ give/ training/ in front/ of a// packed audience.
(c/d) Although Brian is rather reluctant, he will/ give/ . . .

(a) By means of practicing over and over, one can/ develop/ skills/ without/ formal// instruction.
(b) By imitating its mother's voice, a baby can develop/ speech/ without/ formal// instruction.
(c/d) Children in all cultures can/ develop/ . . .

(a) Whenever they leave all the lights on, they/ waste/ energy/ that/ could have// been used much better.
(b) By having detectives write parking tickets, they/ waste/ skills/ that/ could have// been used much better.
(c/d) Since they are not very sensible, they/ waste/ . . .

(a) By taking illegal steroids, some athletes might/ improve/ performance/ after/ a// little while.
(b) Allowing a completely free market might/ improve/ competition/ after/ a// little while.
(c/d) In a way, these new measures might/ improve/ . . .

(a) As the job is complicated to carry out, it will/ require/ planning/ beyond/ what is// already done.
(b) As the report has to be comprehensive, it will/ require/ material/ beyond/ what we// already have.
(c/d) Given the poor state the project is in, it will/ require/ . . .

(a) By increasing productivity quickly, they were able to/ satisfy/ demand/ from/ its// customers.
(b) By setting aside money for lawsuits, they were able to/ satisfy/ claims/ from/ its// customers.
(c/d) By spending a lot of money, that company was able to/ satisfy/ . . .

(a) Since the talks are such a success, it's hard to/ meet/ demand/ for public/ speakers// on various topics.
(b) Since the budget is so tight, it's hard to/ meet/ levels/ for public/ spending// that will satisfy everyone.
(c/d) The meeting in the town hall was about how to/ meet/ . . .

(a) As Ian is now in charge of the law firm, he likes to/ exercise/ power/ by having// other people do the tedious jobs.
(b) As Ian is now the firm's fitness trainer, he likes to/ exercise/ staff/ by having// them running up and down the stairs.
(c/d) Unfortunately, this new manager likes to/ exercise/ . . .

(a) By means of this official ceremony, we/ pay/ tribute/ to the/ veterans abroad.
(b) Instead of giving foreign currency, we/ pay/ dollars/ to the/ veterans abroad.
(c/d) It is not more than reasonable that we/ pay/ . . .

(a) By spelling out discrepancies in the record, the lawyer will/ cast/ doubt/ while/ at// the same time providing an alternative account.
(b) As he needs to recruit potential stars, the director will/ cast/ girls/ while/ at// the film studios in sunny Los Angeles.
(c/d) It might be hard to believe that Martin will/ cast/ . . .

(a) By stressing the risks of obesity, the doctor tried to/ attach/ importance/ to the// values of a healthy lifestyle.
(b) By stating what is and is not allowed, George tried to/ attach/ boundaries/ to the// values of free expression.
(c/d) Believe it or not, but George tried to/ attach/ . . .

(a) As both positions are so far apart, they will never/ reach/ agreement/ acceptable// to most.
(b) Since they don't care about quality, they will never/ reach/ standards/ acceptable// to most.
(c/d) I am totally convinced that they will never/ reach/ . . .

(a) Since the senator has a short temper, he might/ lose/ control/ in an/ instant.
(b) If the unsavory truth comes out, the senator might/ lose/ support/ in an/ instant.
(c/d) We are indeed afraid that the senator might/ lose/ . . .

(a) By making the most of your abilities, you can/ ensure/ success/ in your/ first year// on the job.
(b) By stating your objectives beforehand, you can/ ensure/ results/ in your/ first year// on the job.
(c/d) I am quite confident that you can/ ensure/ . . .

(a) As they think they're on the right track, they do not/ reverse/ direction/ and take// an unexpected path.
(b) As they are extremely principled, they do not/ reverse/ decisions/ and take// an unexpected path.
(c/d) The party decided in the end to not/ reverse/ . . .

(a) As they weren't allowed to look away, the sisters had to/ bear/ witness/ when/ their// father killed their pet.

(*Appendixes continue*)

(b) Because of the horrible environment, the sisters had to/ bear/ tragedy/ when/ their// father killed their pet.
(c/d) When they were very young, the sisters had to/ bear/ . . .

(a) To see whether the tumor is malignant, Luke will/ undergo/ surgery/ in a/ well-// known hospital.
(b) After the blood is drawn from the patient, it will/ undergo/ testing/ in a/ well-// known laboratory.
(c/d) It came as a shock to hear that Tony will/ undergo/ . . .

(a) By acting nonchalantly the spy was able to not/ arouse/ suspicion/ among/ the// different people.
(b) By being friendly, the negotiator was able to not/ arouse/ hostility/ among/ the// different people.
(c/d) In the end, we were able to not/ arouse/ . . .

(a) In the blank fields on the form, you have to/ include/ information/ that/ can be// relevant to your claim.
(b) In the yearly progress report, you have to/ include/ development/ that/ can be// relevant to your bonus pay.
(c/d) It is standard practice that you have to/ include . . .

(a) The office of the Dean of Studies can indeed/ offer/ advice/ to a/ wide range of// questions.
(b) The right political connections can indeed/ offer/ access/ to a/ wide range of// positions.
(c/d) The newly established office can indeed/ offer/ . . .

(a) Having fled their war-torn country, these people/ seek/ refuge/ with/ a lot of// anticipation.
(b) By speaking to the Dalai Lama, these people/ seek/ wisdom/ with/ a lot of// anticipation.
(c/d) I am not at all surprised that these people/ seek/ . . .

(a) In order to blend into the environment, chameleons can/ change/ color/ in a/ very// short time.
(b) Whenever an alpha male is beaten, chimp troupes can/ change/ leader/ in a/ very// short time.
(c/d) Rebecca read in a book that this flock of geese can/ change/ . . .

(a) By using one of his master keys, the janitor could/ gain/ access/ without/ much// trouble.
(b) Being plugged into the outlet, the battery could/ gain/ energy/ without/ much// trouble.
(c/d) The scientists hoped that the robot could/ gain/ . . .

(a) By doing the best that one can is it possible to/ achieve/ results/ that/ will be// satisfactory.
(b) With smart accounting, they think it is possible to/ achieve/ figures/ that/ will be// convincing.
(c/d) As I said before, oftentimes it is possible to/ achieve/ . . .

(a) As they always make touchdowns, the team will/ score/ points/ and might/ ultimately// win the championship.
(b) As their closest opponents lost, the team will/ score/ places/ and might/ ultimately// win the championship.
(c/d) Ron tries to find ways so that his team will/ score/ . . .

(a) Before publishing these documents, we have to/ obtain/ permission/ from/ the// distributor.
(b) Since we're lacking some vital parts, we have to/ obtain/ components/ from/ the// distributor.
(c/d) It is part of the agreement that we have to/ obtain/ . . .

(a) In order to resolve their differences, they will/ hold/ talks/ of any/ importance at// the headquarters.
(b) In order to protect them against theft, we will/ hold/ items/ of any importance at// the headquarters.
(c/d) For obvious security reasons, they will/ hold/ . . .

(a) Good marriage consultants should try to/ resolve/ disputes/ in a/ relatively// impartial way.
(b) People who write dictionaries should try to/ resolve/ meanings/ in a/ relatively// impartial way.
(c/d) The dictionary editor should try to/ resolve/ . . .

(a) By explaining things clearly, the instructor could/ avoid/ confusion/ about/ how to// proceed with the task.
(b) By covering up his tracks, the fugitive could/ avoid/ discovery/ about/ how he// escaped from prison.
(c/d) Although it was not easy, the foreigner could/ avoid/ . . .

(a) Since the movie is a real hit, spin-offs will/ follow/ suit/ whether/ we like it// or not.
(b) As precious metals prices are linked, silver will/ follow/ gold/ whether/ we like it// or not.
(c/d) It is anticipated that the price of silver will/ follow/ . . .

(a) As they are little criminals, they are likely to/ cause/ trouble as soon/ as we// leave the house.
(b) As this wire is almost broken, it is likely to/ cause/ failure/ as soon/ as we// leave the house.
(c/d) The rushed repairs are quite likely to/ cause/ . . .

(a) As stated in his father's will, Marc can/ lay/ claim/ on almost/ everything.
(b) Since he has been a mason for so long, Marc can/ lay/ stone/ on almost/ anything.
(c/d) It is an exaggeration to say that my uncle can/ lay/ . . .

(a) Since you haven't prepared well, you might/ encounter/ problems/ when/ you take the// final exam.
(b) If you go to the games arcade, you might/ encounter/ children/ when/ you get to the// video games.
(c/d) As I told you before, you might/ encounter/ . . .

(a) By handing out leaflets, we will for sure/ raise/ awareness/ and increase/ public// interest.
(b) These bad economic decisions will for sure/ raise/ inflation/ and increase/ interest// rates.
(c/d) Taking such drastic measures will for sure/ raise/ . . .

(a) In order to get enough food ready, Fanny will have to/ spend/ hours/ in a/ hot// kitchen.
(b) As the paper needs to be exhaustive, Fanny will have to/ spend/ weeks/ in a/ hot// library.
(c/d) It has come to the point that Fanny will have to/ spend/ . . .

Appendix B

Experiment 2: Items

For each item, the constraining context/ high-transitional probability (C–H) version appears first (a), followed by the constraining context/ low-transitional probability (C–L) version (b). The (c/d) version is the neutral context, which was the same for the high-transitional probability (N–H) and the low-transitional probability (N–L) versions. The rest of the sentence was the same as in the constraining context versions, starting from

the noun. The / symbols delimit the regions of analysis, the // symbols indicate a line break. The space before the first word in a region of analysis was included in the analysis.

To be comparable to the filler items, we added a second sentence after the critical sentence. As this sentence was the same for all conditions, we only provide it for version (a).

(a) In addition to a general outline, you have to/ include/ details/ that/ are in// any way relevant. But someone must have told you that before.
(b) To calculate a team's ranking, you have to/ include/ results/ that/ are in// any way relevant.
(c/d) It is standard practice that you have to/ include/ . . .

(a) Since the job involves many different parts, it will/ require/ planning/ to an/ even// higher level than I expected. And I had already expected much.
(b) Staying in shape and good health will/ require/ exercise/ to an/ even// higher level than I expected.
(c/d) I was dismayed when Donald told me that it will/ require/ . . .

(a) As we think we're on the right track, we do not/ reverse/ direction/ under/ any// circumstance. Being too stubborn might be bad though.
(b) As they are extremely principled, they do not/ reverse/ decisions/ under/ any// circumstance.
(c/d) The group members are adamant that they do not/ reverse/ . . .

(a) As the product is such a huge success, it is hard to/ satisfy/ demand/ in an// acceptably short time. Everyone had expected that though.
(b) Because of the many lawsuits by customers, it is hard to/ satisfy/ claims/ in an// acceptably short time.
(c/d) By spending a lot of money, that company was able to/ satisfy/ . . .

(a) With the bartenders on strike, they cannot/ serve/ drinks/ even/ if you were to// ask for it. You will have to go somewhere else if you want to get drunk.
(b) As there are minors here, they cannot/ serve/ liquor/ even/ if you were to// ask for it.
(c/d) The notice on the wall says that they cannot/ serve/ . . .

(a) Special passes can be obtained that will/ allow/ access/ without/ having to// turn away too many people. We will have to see whether it works out.
(b) Having several options available will/ allow/ choice/ without/ having to// take whatever is offered.
(c/d) It is rather unclear how they will/ allow/ . . .

(a) By giving tax breaks to businesses, we might/ encourage/ investment/ though/ it// can take a while. Other measures might be necessary as well.
(b) Using propaganda and dissent, the rebels might/ encourage/ opposition/ though/ it// can take a while.
(c/d) Most people think that this legislation might/ encourage/ . . .

(a) In order for the army to function well, soldiers have to/ obey/ orders/ as much/ as// is reasonable. Some kind of discipline is certainly necessary.
(b) In more hierarchical societies, children have to/ obey/ adults/ as much/ as// is reasonable.
(c/d) Some youngsters think it's laughable that they have to/ obey/ . . .

(a) After a traumatic event, many people do not like to/ share/ experiences/ even/ not// with someone they are close with. It takes a long time to overcome this.
(b) Young children are often jealous and do not like to/ share/ possessions/ even/ not// with someone they are close with.
(c/d) It is unfortunate that so many people do not like to/ share/ . . .

(a) Because of delusions or hallucinations, you might/ hear/ voices/ coming/ from// everywhere. We advise you to stay calm.
(b) Since most babies don't like to fly, you might/ hear/ crying/ coming/ from// everywhere.
(c/d) We know it can be very annoying, but you might/ hear/ . . .

(a) In order to blend into the environment, chameleons can/ change/ color/ in a/ very// short time. Not everyone believes that though.
(b) Many people believe that the word of God can/ change/ lives/ in a/ very// short time.
(c/d) Rebecca has told me numerous times that they can/ change/ . . .

(a) Since you are ultimately responsible, you should/ exercise/ control/ as much/ as// possible. Good judgment will be useful.
(b) Since these terms can be insulting, you should/ exercise/ caution/ as much/ as// possible.
(c/d) For your own peace of mind, you should/ exercise/ . . .

(a) Because of procedural error, the witness could no longer/ give/ evidence/ even/ if/ he wanted to. We had never expected that.
(b) Since it is clearly our own fault, we can no longer/ give/ excuses/ even/ if// we wanted to.
(c/d) We were all puzzled to hear that Fiona can no longer/ give/ . . .

(a) By stressing the risks of obesity, the doctor tried to/ attach/ importance/ to the// issue at hand. It should certainly be taken seriously.
(b) With these restrictions on custody, Ann tried to/ attach/ boundaries/ to the// issue at hand.
(c/d) Believe it or not, but George tried to/ attach/ . . .

(a) Illegal steroids are thought to/ improve/ performance/ after/ a little while.// But the side effects might be devastating.
(b) Curbing monopolies is thought to/ improve/ competition/ after/ a little while.//
(c/d) These new measures are thought to/ improve/ . . .

(a) In order to improve health care, senators plan to/ introduce/ legislation/ as soon// as possible. However, even that can take a while.
(b) So the jury can see the wreckage, lawyers plan to/ introduce/ photographs/ as soon// as possible.
(c/d) It is a rumor that these gentlemen plan to/ introduce/ . . .

(a) If you want to invest, a bank can usually/ offer/ advice/ free/ of any// extra charges. I think it's always good to check it out thoroughly.
(b) To improve flexibility, investment plans usually/ offer/ choice/ free/ of any// extra charges.
(c/d) It was a relief to hear that institutions usually/ offer/ . . .

(a) As she has to fill in many pages, Fanny will/ spend/ hours/ in the/ university// library. Her boyfriend is already griping about it.
(b) Instead of only a few days, Fanny will/ spend/ weeks/ in the/ university// library.
(c/d) It has come to the point that Fanny will/ spend/ . . .

(a) Since Jill did not study well, she will/ make/ mistakes/ without/ even being// aware of it. And don't even try telling her.
(b) In order to underline her point, she will/ make/ gestures/ without/ even being// aware of it.
(c/d) Everyone expects it now that Jill will/ make/ . . .

(a) By providing child care, the council can/ help/ parents/ who are/ having// a hard time. We agree that this is a worthy issue.

(*Appendixes continue*)

(b) Many believe that a strong democratic union can/ help/ workers/ who are/ having// a hard time.
(c/d) Since George is always encouraging, he can/ help/ . . .

(a) As the federal reserve cut interest rates, banks will/ follow/ suit/ whether/ we// like it or not. Not that it will have much impact on me though.
(b) As precious metals prices are linked, silver will/ follow/ gold/ whether/ we// like it or not.
(c/d) It is anticipated that this material will/ follow/ . . .

(a) Using appropriate protection and handling might/ prevent/ damage/ though/ it is// not a foolproof remedy. Maybe other measures will be better.
(b) To some, the fear of capital punishment might/ prevent/ murder/ though/ it is// not a foolproof remedy.
(c/d) Implementing these regulations might/ prevent/ . . .

(a) According to these terms, the company will/ supply/ goods/ but will/ not take any// further responsibilities. In other words, liability is restricted.
(b) In addition to clothes, the company will/ supply/ shoes/ but will/ not take any// further responsibilities.
(c/d) Since we made the commitment, we will/ supply/ . . .

(a) As I just bought a fancy cell phone, I can now/ send/ messages/ though/ they// cannot be too long. You can say a lot in a few words though.
(b) As the magazine invites contributions, you can now/ send/ articles/ though/ they// cannot be too long.
(c/d) The manager told you before that you should/ keep/ . . .

(a) To keep their hands warm in winter, they like to/ wear/ gloves/ at least/ as// long as they are well made. The price is of lesser importance.
(b) Instead of dark colors, doctors and nurses like to/ wear/ whites/ at least/ as// long as they are well made.
(c/d) Brett has often wondered why so many women like to/ wear/ . . .

(a) In general, people expect that local authorities will/ provide/ services promptly// upon request. But we're afraid the bureaucracy can be awful.
(b) In general, most car insurance policies will/ provide/ coverage/ promptly// upon request.
(c/d) We are very delighted to hear that the company will/ provide/ . . .

(a) With lots of publicity, the small museum was able to/ attract/ visitors/ from a// diverse range of groups. I was surprised at that.
(b) More than attention, the weeping child was able to/ attract/ sympathy/ from a// diverse range of people.
(c/d) It is fun to see how easily they can/ attract/ . . .

(a) As Brian has a short temper, he will surely/ lose/ control/ in a/ short// while. It might be necessary to seek some assistance.
(b) As it drew much opposition, the measure will surely/ lose/ support/ in a/ short// while.
(c/d) We are indeed afraid that the senator might/ lose/ . . .

(a) Through weekly reports and intense testing, we will/ monitor/ progress/ so that// nobody will get upset. The task will surely be tedious.
(b) Since the town has huge debts, the comptroller will/ monitor/ spending/ so that// nobody will get upset.
(c/d) From tomorrow on, the man with the clipboard will/ monitor/ . . .

(a) Some people write to agony columns because they want to/ seek/ advice/ no matter// what. Whether it will help in the long term is debatable.
(b) Not happy with the current status quo, we want to/ seek/ change/ no matter// what.
(c/d) I am not at all surprised that these people want to/ seek/ . . .

(a) In addition to freight, these boats also always/ carry/ passengers/ even/ though// some people have objected to it. You can never make everyone happy.
(b) As they are addicted to nicotine, they always/ carry/ cigarettes/ even/ though// some people have objected to it.
(c/d) Since it has been a custom for long, they always/ carry/ . . .

(a) Unable to get to the victims, the rescuers had to/ bear/ witness/ when/ the rest// of the family perished. It was truly heartbreaking.
(b) Because their misfortune was great, they had to/ bear/ tragedy/ when/ the rest// of the family perished.
(c/d) When they were very young, the sisters had to/ bear/ . . .

(a) Since Fred is the man in charge, he has to/ take/ decisions/ even/ though he// would rather go home. He is just not feeling too well.
(b) After the candidate has given his talk, he has to/ take/ questions/ even/ though he// would rather go home.
(c/d) While Robert might not want to, he has to/ take/ . . .

(a) Faced with inaccuracies, copy editors will/ correct/ errors/ even/ without// asking. They claim it is part of their job.
(b) In order to make them more glossy, magazines will/ correct/ photos/ even/ without// asking.
(c/d) Lisa was upset to hear that they will/ correct/ . . .

(a) New equipment and division of labor could/ increase/ efficiency/ in a/ certain// way. In any case, I don't expect too much too soon.
(b) Threatening unruly kids with grounding could/ increase/ discipline/ in a/ certain// way.
(c/d) It is hard to believe that these things could/ increase/ . . .

(a) People think that subliminal messages can/ influence/ behavior/ and some/ people// don't like that. We don't think this will be solved soon.
(b) Restrictions on what can be built where can/ influence/ planning/ and some/ people// don't like that.
(c/d) It is not always clear what exactly can/ influence/ . . .

(a) Terrorists have stated that they can/ launch/ attacks/ from/ almost any// place in the world. Maybe I should move to Mars.
(b) Nowadays, military submarines can/ launch/ rockets/ from/ almost any// place in the world.
(c/d) It is disconcerting that these people can/ launch/ . . .

(a) By setting strict limits on hunting we try to/ protect/ wildlife/ as much/ as// possible. It will never be enough though.
(b) After an attack on a holiday resort, they try to/ protect/ tourists/ as much/ as// possible.
(c/d) Aware of their reputation, these states try to/ protect/ . . .

(a) Thanksgiving is always a good time if one wants to/ visit/ relatives/ who are/ on// their own. It might be most rewarding if one doesn't go alone.
(b) The warden said weekends are best if one wants to/ visit/ prisoners/ who are on// their own.
(c/d) It can sometimes be a hardship if one wants to/ visit/ . . .

(a) As Ian can prove the impounded bike is his, he will/ lay/ claim/ on what/ looks// to be his vehicle. Miriam told me of this fact.

(b) As it is part of his job as a mason, Marc will/ lay/ brick/ on what/ looks// to be a shabby wall.
(c/d) It is mere speculation to say that they will/ lay/ . . .

(a) In order to cross the river, the engineers will/ build/ bridges/ which/ will// look good as well. In the end, everyone will be happy.
(b) In order to exhibit paintings and art, cities will/ build/ museums/ which/ will// look good as well.
(c/d) After quite a long time planning, they will/ build/ . . .

(a) In order to estimate attendance, we should/ keep/ records/ for at/ least// two years. Some people even think that that's too short.
(b) In order to build a better team, you should/ keep/ players/ for at/ least// two years.
(c/d) Although it took a little while, you can now/ send/ . . .

(a) Since they always want to win, the team won't/ accept/ defeat/ even/ when they// are far behind. Unfortunately they have to do that often.
(b) Because they need a lot of wins, the team won't/ accept/ losses/ even/ when they// are far behind.
(c/d) It is silly that they simply won't/ accept/ . . .

(a) By explaining things clearly, the teacher tries to/ avoid/ confusion/ whenever/ he// can. Luckily for him, he almost always succeeds.
(b) As he is allergic to clams and the like, Ed tries to/ avoid/ shellfish/ whenever/ he// can.
(c/d) Although it is not easy, the foreigner tries to/ avoid/ . . .

(a) By means of practicing the same thing a lot, people can/ develop/ skills/ all on// their own. It seems as if repetition aids learning.
(b) By imitation and using their vocal tracts, babies can/ develop/ speech/ all on// their own.
(c/d) Apparently, children in all cultures can/ develop/ . . .

(a) By making the most of your abilities, you can/ ensure/ success/ in almost/ anything// that you want to do. But ultimately it is all up to you.
(b) By stating the principles beforehand, you can/ ensure/ clarity/ in almost/ anything// that you want to do.
(c/d) I am very much convinced that you can/ ensure/ . . .

(a) By using one of his master keys, the janitor could/ gain/ access/ without/ much// trouble. What happened next is still a mystery though.
(b) Being plugged into the outlet, the battery could/ gain/ energy/ without/ much// trouble.
(c/d) The scientists hoped that the robot could/ gain/ . . .

(a) Only the committee on gambling is in the position to/ grant/ licenses/ to the// people who apply for it. We expect many applications.
(b) By using an official waiver, we are in the position to/ grant/ immunity/ to the// people who apply for it.
(c/d) Following today's decision, we are in the position to/ grant/ . . .

(a) By increasing the production in a major way, we will/ meet/ demand/ when/ it is// necessary. It is always good to be prepared.
(b) As the climb is tricky and fog is coming in, you will/ meet/ danger/ when/ it is// unexpected.
(c/d) The meeting in the hall is about how we will/ meet/ . . .

(a) Surrounded by DEA agents, Jeffrey could not/ resist/ arrest/ and quickly/ gave// up. I guess it's in his character.
(b) As Jeffrey loves dairy products, he could not/ resist/ cheese/ and quickly/ gave// in.
(c/d) As his partner had predicted, Jeffrey could not/ resist/ . . .

(a) The United Nations was established in order to try to/ resolve/ disputes/ in a// civilized way. At least, that is the intention.
(b) Our customers service is set up in order to try to/ resolve/ mistakes/ in a// civilized way.
(c/d) While it is an arduous task, the group will try to/ resolve/ . . .

(a) By leaving all the lights on, they surely/ waste/ energy/ in a/ very// irresponsible way. I would have acted differently for sure.
(b) By keeping the best players on the bench, they surely/ waste/ talent/ in a/ very// irresponsible way.
(c/d) Since they are not acting very sensible, they surely/ waste/ . . .

(a) Some claim that raising the minimum wage will/ destroy/ jobs/ but most/ people// don't seem to care. This carelessness might hurt them in the long run.
(b) As the slaughterhouse is filthy, inspectors will/ destroy/ meat/ but most/ people// don't seem to care.
(c/d) Bill was outraged when he heard that they will/ destroy/ . . .

(a) As you signed the contract, you are eligible to/ receive/ payment/ without/ any// more problems. It used to be much harder to get.
(b) Even with bad credit, car buyers are eligible to/ receive/ finance/ without/ any// more problems.
(c/d) Jon is happy to hear that at last he is eligible to/ receive/ . . .

(a) Israelis and Arabs are doubtful that the new plan will/ bring/ peace/ any time/ soon// now. Unfortunately the wait has been very long already.
(b) As an accompaniment to the soup, the server will/ bring/ bread/ any time/ soon// now.
(c/d) Margaret is not very confident that they will/ bring/ . . .

(a) Instead of quick emails, you can also/ write/ letters/ as long/ as// you are motivated to do so. But motivation seems to be lacking a bit.
(b) In addition to some letters, my child can also/ write/ numbers/ as long/ as// he is motivated to do so.
(c/d) I was pleasantly surprised to see that Sean can also/ write/ . . .

Received August 24, 2004
Revision received February 14, 2005
Accepted February 22, 2005 ■