Journal of Psycholinguistic Research, Vol. 32, No. 1, January 2003 (© 2003)

Integration of Syntactic and Semantic Information in Predictive Processing: Cross-Linguistic Evidence from German and English

Yuki Kamide,[1,3] Christoph Scheepers,[2] and Gerry T. M. Altmann[1]

Two visual-world eyetracking experiments were conducted to investigate whether, how, and when syntactic and semantic constraints are integrated and used to predict properties of subsequent input. Experiment 1 contrasted auditory German constructions such as, "The hare-nominative eats . . . (the cabbage-acc)" versus "The hare-accusative eats . . . (the fox-nom)," presented with a picture containing a hare, fox, cabbage, and distractor. We found that the probabilities of the eye movements to the cabbage and fox before the onset of NP2 were modulated by the case-marking of NP1, indicating that the case-marking (syntactic) information and verbs' semantic constraints are integrated rapidly enough to predict the most plausible NP2 in the scene. Using English versions of the same stimuli in active/passive voice (Experiment 2), we replicated the same effect, but at a slightly earlier position in the sentence. We discuss the discrepancies in the two Germanic languages in terms of the ease of integrating information across, or within, constituents.

KEY WORDS: sentence processing; anticipatory eye movements; visual-world paradigm; prediction; Germanic languages.

The order of the authors is arbitrary (all three authors contributed to this study equally). YK is now at the Department of Psychology, University of Manchester. We are grateful to Matthew Crocker who helped us to have this opportunity for cross-linguistic research. Also we would like to thank Pia Knöferle for helping conduct Experiment 1. The present research was supported by a grant from the Medical Research Council, UK (G0000224) (awarded to GTMA) and a grant from Deutsche Forschungsgemeinschaft, Germany (SFB 378-Alpha) (awarded to Matthew Crocker). All the items used in Experiments 1 and 2 can be viewed at www.york.ac.uk/res/prg/jpr02

[1] Department of Psychology, University of York, UK.

[2] Department of Computational Linguistics, Saarland University, Germany.

[3] To whom all correspondence should be addressed: Department of Psychology, University of Manchester, Oxford Road, Manchester M13 9PL, UK. email: yuki.kamide@man.ac.uk

0090-6905/03/0100-0037/0

INTRODUCTION

In recent years, there has been increasing research on the nature of incrementality in human sentence processing (e.g., Frazier, 1979; Marlsen-Wilson, 1975; Sedivy *et al.,* 1999; Tanenhaus *et al.,* 1990). The majority of the research suggests that the human parser incorporates incoming linguistic input into the existing structure, often without any difficulty, as soon as it is encountered within the sentence; the preceding lexical items usually provide enough information for the parser to attach the current item into the existing structure without delay. In this paper, we investigate a feature of incremental sentence processing whereby preceding lexical material enables the partial processing of forthcoming material even before that forthcoming material is encountered within the sentence (cf. Altmann & Kamide, 1999). The aim of this paper is to explore the time-course of such predictive processing in sentence comprehension. Specifically, we look at what types of information, or constraints, can enable prediction—that is, what types of information can be used, and when, to narrow down the range of options for what might be referred to next.

To address these questions, we adopt the visual-world paradigm, in which eye movements around a visual scene are monitored as participants listen to a description of that scene or to an instruction to manipulate objects within the scene (initially Cooper, 1974; followed by Allopenna *et al.,* 1998; Altmann & Kamide, 1999; Chambers *et al.,* 2002; Eberhard *et al.,* 1995; Sedivy *et al.,* 1999; Tanenhaus *et al.,* 2001; Tanenhaus *et al.,* 1995; Trueswell *et al.,* 1999). One important feature of the technique is that each saccadic eye movement can be time-locked to the concurrent auditory stimulus. The paradigm is based on the finding that a saccadic eye movement is always accompanied (in fact, preceded) by an attention shift toward the object that is the destination of the eye movement (Duebel & Schneider, 1996; Hoffman & Surbramaniam, 1995; Kowler *et al.,* 1995; Rayner *et al.,* 1978; Remington, 1980). Thus, the researcher can establish which part of the auditory input triggers the attention shift. Tanenhaus *et al.* (1995) showed that the visual-world methodology offers experimental environments in which visual contexts serve in the same way as preceding sentential contexts as conventionally used in reading experiments (cf. Altmann & Steedman, 1998).

The experiments in this paper adopt the version of the visual-world paradigm used by Altmann and Kamide (1999): Instead of real objects, the visual stimuli were semirealistic scenes presented on a computer screen, and the participants were asked simply to “look and listen.” In that study, Altmann and Kamide established that verbs’ *selectional restrictions* (semantic constraints on permissible arguments of verbs) can be used to predict the forthcoming direct object even before the onset of that referring expression.

For example, on hearing "The boy will eat . . .," participants looked toward the one edible object in the scene more often than on hearing "The boy will move . . ."; there were several moveable objects in the scene, but there was just one edible object, and thus the selectional restrictions associated with "eat" could narrow down the domain of subsequent reference more fully than could those associated with "move." Following Altmann & Kamide (1999), we shall refer to such saccadic eye movements as *anticipatory*.

The majority of studies on anticipatory eye movements have revealed the use of verb information in prediction of postverbal arguments in head-initial structures (e.g., Altmann & Kamide, 1999; Kako & Trueswell, 2000; Kamide *et al.*, in press—Experiments 1 & 2; Sussman & Sevidy, 2001). These findings offer a useful addition to previous evidence that verb information determines or modulates initial attachments of postverbal arguments as those arguments are encountered (e.g., MacDonald *et al.*, 1994; Trueswell *et al.*, 1993; but see also Mitchell, 1987); unlike these earlier results, the recent findings on prediction suggest that verbs' constraints on what kinds of argument can follow modulate the partial interpretation of the sentence constructed thus far as soon as the verb is encountered and not simply when those arguments are encountered. This evidence for an immediate effect of the verb is compatible with the hypothesis that information stored with the lexical head drives the processing of the nonhead constituents (cf. head-driven parsing: Pritchett, 1991, 1992; HPSG: Pollard & Sag, 1987, 1994). However, heads may provide useful sources of predictive information only insofar as they precede their arguments (in languages like English); in a language in which heads follow their arguments, other sources of information might be available on which basis to predict upcoming information. Kamide *et al.* (in press, Experiment 3) have addressed this question using the head-final language Japanese. The main aim of their study was to investigate whether prehead constituents can drive prediction. They presented participants with a picture containing, for example, a waitress, a customer, a hamburger, and a dustbin. Simultaneously, participants heard either "waitress-nominative customer-dative merrily hamburger-accusative give." ("weitoresu-ga kyaku-ni tanosigeni hanbaagaa-o hakobu" in Japanese; "The waitress will merrily bring the hamburger to the customer" in English) or "waitress-nominative customer-accusative merrily tease." ("weitoresu-ga kyaku-o tanosigeni karakau" in Japanese; "The waitress will merrily tease the customer" in English). The crucial difference between the two conditions is the case-marker of the second argument. The manipulation is based on the property of Japanese grammar that an accusative-marked noun phrase (typically serving as the Theme) is required after a sequence consisting of a nominative noun phrase and a dative noun phrase (unless the verb is one of the small set of dative-taking ditransitive verbs). In the example scene, the hamburger is the most plausible Theme in the situation where the waitress is the Agent and the customer is the Goal (in the customer-dative condition). In contrast, in the customer-accusative condition, the waitress is again the Agent but the customer is now the Theme, and although an upcoming Goal is licensed by the grammar, the hamburger is unlikely to be mentioned in Goal position. Thus, it was predicted that the hamburger should be looked at more often after "waitress-nominative customer-dative" than after "waitress-nominative customer-accusative." Precisely such a pattern was observed (during the adverb "merrily"), indicating that the case-marking information combined with real-world plausibility information to enable the prediction of the third argument. This evidence for prehead prediction is compatible with previous experiments that have demonstrated prehead attachment in the processing of head-final structures (e.g., Bader & Lasser, 1994; Kamide & Mitchell, 1999; Koh, 1997) and is incompatible with accounts of parsing that are strictly head-driven (Pritchett, 1991, 1992).

The Kamide *et al.* (submitted) data on Japanese demonstrate that a sequence of noun phrase (NP1-NP2 can predict a subsequent NP3 even when the verb will appear sentence finally. The case-marking associated with NP1 and NP2, combined with real-world knowledge, enables the prediction. In another study, Kamide *et al.* (submitted, Experiment 2) demonstrated that in English, the sequence NP1-V can predict a subsequent NP2 (in an NP1-V-NP2 structure) on the basis of the combination of thematic information found in NP1 and selectional restrictions from the verb. In that study, participants were presented with a visual scene portraying a man, a young girl, a motorbike, and a fairground carousel (among other things—a pint of beer, a jar of sweets). Participants heard, simultaneously, either "The man will ride the motorbike" (the most plausible thing for the man to ride) or "The girl will ride the carousel" (the most plausible thing for the girl to ride). During the verb "ride," more looks were observed to the plausible object than to the implausible object (that is, more looks to the motorbike following "the man," and to the carousel following "the girl"). Importantly, this pattern was not found when a different verb was used (one that selected for other objects in the scene—for this example scene, "taste")—ruling out a direct association between "the man" and the motorbike, or "the girl" and the carousel.

In the Japanese study described earlier (Kamide *et al.*, in press, Experiment 3), prediction of an upcoming argument was based on the combination of morphosyntactic information derived from two constituents of the same type (both noun phrases). In the English study just described (Kamide *et al.*, in press, Experiment 2), on the other hand, prediction of an upcoming argument was based on the *combination* of semantic information derived from a verb and the noun phrase that preceded it. In the first of two experiments we report here, we explore whether semantic information derived at the verb can combine with morphosyntactic marking on the noun phrase that precedes it. Thus, we address the question of whether and how syntactic information extracted from constituents in one syntactic category can interact with other types of information, such as semantic or pragmatic information, derived from constituents in another syntactic category. In Japanese, for example, case-marking information might interact only with noun-extracted information, because, at least in Japanese, case-markers are restricted to noun phrases. The first experiment described here will explore a situation in which the combination of an explicitly case-marked noun phrase with semantic information associated with the subsequent verb might predict a postverbal noun phrase. The experiment uses main clause NP1-V-NP2 constructions in German.

According to German morphology, masculine singular noun phrases (but not only those) are unambiguously marked for case (e.g., "Der Kater aß den Fisch"/ The tomcat-nominative ate the fish-accusative). The case-marking information is vital for analyzing the syntactic function and/or the thematic role of an argument, because German, unlike English, allows variable constituent ordering: A variant of the example above—namely, "Den Fisch aß der Kater" (The fish-accusative ate the tomcat-nominative)—describes the same event (the tomcat being the eater, and the fish being eaten), just in a reverse ordering of subject and object. This contrasts with English where the order of syntactic functions is fixed.

The main experiment (Experiment 1) in this paper takes advantage of variable function ordering in German to investigate prediction. Using the visual-world eyetracking paradigm mentioned above, the experiment consists of two sentential conditions as follows:

(1) a. Der Hase frißt gleich den Kohl.
The hare-nom eats shortly the cabbage-acc.
"The hare will shortly eat the cabbage."
b. Den Hasen frißt gleich der Fuchs.
The hare-acc eats shortly the fox-nom.
"The fox will shortly eat the hare."

Auditory versions of these sentences were presented with a visual scene portraying, for this example, a hare, a cabbage, a fox, and a tree (distractor) (Fig.1).

The question we are interested in is whether the second noun phrase is predicted before the onset of the referring noun. The prediction can be achieved if the verb and the case markers provide sufficient information to determine which object in the accompanying picture is most likely to play the remaining thematic role in the argument structure of the verb. In (1a)—hereafter the *nominative* condition—"Der Hase frißt" (The hare-nominative eats) suggests that the hare is the Agent (subject). Given this, and the action

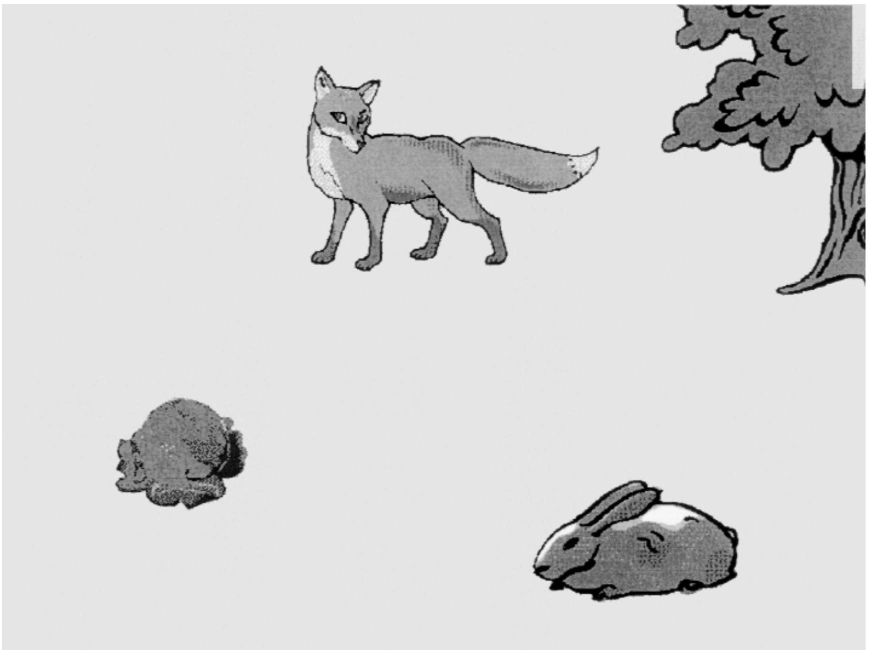

Fig. 1. An example of the visual stimuli used in Experiments 1 and 2.

denoted by the verb (eat), the cabbage might be predicted to be the Theme (direct object), as it is the most plausible object for the hare to eat in the concurrent scene. In contrast, in (1b)—hereafter the *accusative* condition—"Den Hasen frißt" (The hare-accusative eats) indicates that the hare is likely to serve as the Theme of the eating. This might trigger the prediction that the fox should follow, because it is the most likely object in the concurrent scene to eat the hare. In eyetracking terms, anticipatory eye movements toward the "cabbage" objects in the nominative condition (i.e., the objects across the different items that correspond to the cabbage in this example) and toward the "fox" objects in the accusative condition should be obtained prior to the onset of the postverbal noun phrase. Such a pattern of anticipatory eye movements would imply that case-marking information on a noun phrase, along with the semantic constraints associated with a verb, can together combine with real-world knowledge to predict what will most plausibly be referred to in the postverbal position.

Experiment 1 below offers an opportunity to investigate when morphosyntactic information from noun phrases is integrated with a verb's semantic constraints. In Experiment 2 below, conducted in English, we explore how morphosyntactic marking on the verb is integrated with that verb's semantic constraints. It uses the same scenes from Experiment 1, but with sentences such as the following:

(2) a. The hare will eat the cabbage.
    b. The hare will be eaten by the fox.

Example (2a) is the English equivalent of (1a), while (2b) is the passive version of (1b). Thus, the appropriate object to be predicted in (2a)—hereafter the *active* condition—is the cabbage, whereas it is the fox in (2b)—hereafter the *passive* condition. As in the German pair, the first noun phrase in (2a) and (2b) contains the same noun. However, unlike in the German sentences, the thematic role to be played by the first noun does not become certain until the verb is encountered. In principle, the thematic role played by the first noun becomes clear at the noun itself, given its case marking. Thus, the thematic role assignment appropriate for the first noun phrase resolves later in English than in German. Indeed, if one went ahead with role assignment before the verb using probabilistic information about the typical thematic role of the first noun in English (presumably the Agent; cf., Bever, 1970), it would turn out to be incorrect in the passive condition (in German there is an equivalent agent-first preference when the NP1's case is ambiguous between nominative and accusative; see Hemforth & Konieczny, 2000). Consequently, in the English active case, the verb either causes—or confirms—the assignment of the Agent to NP1. In the German nominative condition, no such thematic processing is required at the verb. Conversely, in the English passive, the verb either causes the assignment of Theme to NP1 or causes the *reassignment* of a role to NP1 (in the case where the Agent role was provisionally assigned on the basis of an agent-first preference). Experiment 2, and its comparison with Experiment 1, allows us to investigate the consequences for predictive processing of the differences in the time-course of thematic role assignment in the two languages.

EXPERIMENT 1

As mentioned, the major purpose of Experiment 1 is to explore the processes by which case-marking information (syntactic information) is integrated with verbs' semantic constraints on their arguments in prediction in German sentence processing, using the visual-world eye-tracking paradigm.

Method

Participants

Forty subjects from the Saarland University student community were paid for their participation in the study. All were native speakers of German and either had uncorrected vision or wore soft contact lenses or eyeglasses.

Materials

There were 16 experimental pictures (e.g., Fig. 1), each paired with two sentential conditions (e.g., (1a) and (1b)). The pre- and postverbal noun phrases were all masculine (except for one of the trials in which the postverbal noun phrase was feminine), and all were unambiguously marked for case. The visual scenes were created using commercially available ClipArt packages. The scenes were initially constructed using a 16-color palette at a resolution of 640 × 480 pixels, then later converted to 800 × 600 pixel resolution for presentation. Each scene contained four objects: the first noun object (hare), the nominative second noun object (cabbage), the accusative second noun object (fox), and a distractor (tree). The experimental materials can be viewed at www.york.ac.uk/res/prg/jpr02. Two lists of stimuli were created containing each of the 16 experimental pictures but just one version of each sentence pair—8 of the 16 sentences were drawn from the nominative condition, and 8 from the accusative condition. A further 16 items were added as fillers. These were verb-final structures with ditransitive verbs in the subordinate clause, and the subject noun phrases in both matrix and subordinate clauses had overt nominative case. Each filler picture contained five objects creating a semirealistic scene. The sentences were recorded by a male native speaker of standard German (CS). The sound files were presented to participants via two loudspeakers positioned either side of the viewing monitor. The onsets and/or offsets of critical words in the stimulus sentences were marked using a sound editing package for later analysis.

Procedure

Participants were seated in front of a 21″ color display with their eyes approximately 25″ from the display. They wore an SMI EyeLink head-mounted eyetracker, sampling at 250 Hz with a spatial resolution of less than 0.01°. The viewing was binocular, but only the participant's dominant eye was tracked (the right eye for 75% of all participants, as determined by a simple parallax test prior to the experiment). The stimulus display ran at 160 Hz refresh rate in 800 × 600 pixel resolution. The picture stimuli appeared within a frame of 60 pixels from the horizontal edges and 80 pixels from the vertical edges of the screen, covering 80% of the area around the screen's center.

Participants were told that they would see a series of pictures accompanied by auditory stimuli that were related to the context of the pictures. Their task was simply to look at the pictures and listen to the sentences while their eye movements were monitored. Participants were further told that, after some trials, they would see a simple yes/no comprehension question on the screen (e.g., "Wird der Hase gleich wegrennen?" Will the hare shortly run away?), and that their task was to answer these questions by pressing either the left ("yes") or right ("no") arrow button of a standard PC keyboard in front of them. The question was always about the trial that preceded it. Altogether, eight trials were followed by a question (four experimental items, four fillers), and half of the questions suggested a "yes" response, and the other half "no".

Each trial started with the presentation of a fixation cross in the center of the screen. The participant fixated it so that an automatic drift correction could be performed (the trial would not proceed until the cross was fixated). The experimenter then triggered the presentation of the relevant visual and auditory stimuli. The picture appeared, and after a fixed time delay of 1200 ms from picture onset, the sentence was played over the speakers. The sounds typically ended about 2 seconds before the end of the corresponding picture presentation. Experimental pictures were presented for 6500 ms, and filler pictures for 7000 ms (the latter were paired with longer sentences). When the picture disappeared, the next trial was initiated or, alternatively, a comprehension question was presented and the next trial was initiated after the participant's response. After every fourth trial, the eyetracker was recalibrated using a 9-point fixation stimulus. The EyeLink software automatically validated calibrations and the experimenter could, if required, repeat the calibration process if validation was poor. Calibration took approximately 20 seconds. There was a short practice session before the experimental session. The entire experiment lasted approximately 20 minutes.

Results

Participants answered 95% of the questions correctly. We adopted the same procedure for analyzing the eye-movement data generated by the EyeLink system as described in Altmann and Kamide (1999). The utterances were divided into four different regions for analysis purposes, as shown in Fig. 2. Table I shows the percent of trials where there was at least one fixation onto the "cabbage" objects or the "fox" objects in each sentential condition during the NP1, Verb, and Adverb regions.[4,5]

[4] To be precise, the critical region could be extended to the onset of the head noun, not the determiner, in the second noun phrase, as the determiner did not disambiguate the forthcoming head noun in 31 sentences out of all 32 (there was only one sentence with a feminine "cabbage" object whose determiner helped the object singled out). However, to be conservative, we shall not include eye movements that took place in this determiner region in our analyses.

[5] For simplicity, throughout this paper, we report data only in terms of the percentages of trials in which there was at least one fixation onto certain objects. However, we do not mean to imply that this measure is superior to other commonly used analysis measures, such as the numbers of looks or average duration of looks. Overall, our own analyses suggested very similar patterns of results across a range of different measures (see also Scheepers *et al.*, in prep.).

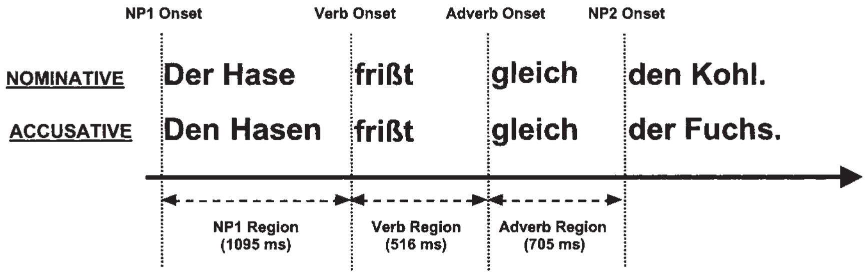

Fig. 2. The region division used for the data analyses in Experiment 1. The numbers in the brackets show the mean duration of the region in milliseconds. The relative sizes of the space between the words in Fig. 2 do not correspond to the actual relative lengths of the pauses in the auditory materials.

An analysis of variance (ANOVA) of the NP1 region data, with Object ("cabbage," "fox") and Case (nominative, accusative) as factors was performed to see whether the anticipatory eye movements were obtained as early as the NP1 region (theoretically, the earliest region at which any prediction might be made). No significant effect was obtained in either the ANOVA or in planned comparisons for the data up to the onset of the verb, suggesting the case-marked noun phrases by themselves did not provide enough information on which to predict the most plausible second noun. For the Verb region, an ANOVA revealed a significant main effect of Object (more looks to the "cabbage" objects than the "fox" objects; $F1(1,39) = 51.93$, $p < .001$; $F2(1,15) = 67.95$, $p < .001$) but no main effect of Case ($F1,F2 < 1$). The interaction was found to be nonsignificant ($F1,F2 < 1$). Planned comparisons suggested that there was no statistical difference between the conditions for

Table I. Percent of Trials with a Fixation onto the "Cabbage" Objects and the "Fox" Objects Obtained in Each Condition for Each of the Three Regions in Experiment 1. The "Cabbage" Objects were the "Appropriate" Objects in the Nominative Condition, and the "Fox" Ones for the Accusative Condition

Region	Condition	"Cabbage" objects	"Fox" objects
NPI region	Nominative	46	46
	Accusative	43	46
Verb region	Nominative	32	15
	Accusative	30	15
Adverb region	Nominative	43	26
	Accusative	38	38

either types of object. Thus the data showed no evidence for anticipatory eye movements before or during the Verb region. However, the data for the Adverb region showed a different pattern. The Object × Case ANOVAs yielded a main effect of Object (more looks to the "cabbage" objects than the "fox" objects; $F1(1,39) = 8.79$, $p < .01$; $F2(1,15) = 6.39$, $p < .025$) and a main effect of Case reliable only by items (accusative > nominative; $F1(1,39) = 1.60$, $p > .10$; $F2(1,15) = 4.74$, $p < .05$). Most crucially, the interaction between Case and Object was significant for the data from the Adverb region ($F1(1,39) = 8.71$, $p < .01$; $F2(1,15) = 7.84$, $p < .02$). This suggests that the *"appropriate"* objects (the cabbage in the nominative condition, the fox in the accusative condition) attracted more fixations than the *"inappropriate"* objects (the fox in the nominative condition, the cabbage in the accusative condition) in this region. Planned comparisons confirmed the general tendency of anticipatory eye movements in the right direction: The "fox" objects attracted statistically more looks in the accusative condition than in the nominative condition ($F1(1,39) = 8.20$, $p < .01$; $F2(1,15) = 8.24$, $p < .02$). The "cabbage" objects were looked at numerically more often in the nominative condition than in the accusative condition, but this difference was not statistically significant ($F1(1,39) = 1.72$, $p > .10$; $F2(1,15) = 1.19$, $p > .10$). We conducted additional ANOVAs on the data from the Adverb region to establish where in the region the interaction between Case and Object became significant by both subjects and items. The analyses revealed that it was approximately 600 ms after the onset of the adverb (approximately 1100 ms after the onset of the verb).

Discussion

The data suggest that there do exist conditions when the most plausible object to be referred to postverbally can be predicted prior to the onset of its corresponding referring expression. This is evidenced by the significantly greater proportion of anticipatory looks toward the "fox" objects in the accusative condition than in the nominative condition. This prediction was achieved by a combination of the precise thematic role assigned to the first noun phrase (on the basis of its case-marking) and the verb's semantic constraints. Together, these combined with real-world knowledge to enable the prediction of what would most likely be referred to postverbally. Experiment 1 provides evidence that, for the purposes of prediction, morphosyntactic information derived from noun phrases can be quickly integrated with semantic information extracted from the verb. The role of the verb here is inferred on the basis that the "appropriate" objects were appropriate only in terms of their compatibility with verbs' semantic constraints; they might not be the most plausible objects to follow if different verbs were used (e.g., "The hare-nom attracts. . ." might prefer the fox than the cabbage as the next noun).[6]

Despite the significant interaction between Case and Object, planned comparisons revealed no statistical effect of case in looks towards the "cabbage" objects (or the corresponding objects across the different trials). It looks as if the lack of effect is due to a relatively large number of fixations in the accusative condition. One possible cause of this might be that the "cabbage" objects had an overall visual advantage over the "fox" ones, presumably because they were either visually salient in some ways (e.g., color, size) or because the NPI objects (e.g., the hare) were facing toward the "cabbage" objects in nearly half of the experimental items. We believe that this could also reflect (and most probably does reflect) the well-documented finding in German that people tend to analyze the clause-initial noun phrase as nominatively marked when case is ambiguous between nominative and accusative (see, e.g., Hemforth & Konieczny, 2000, for a collection of articles on that issue.). Although the first noun phrases in our sentences were not case-ambiguous, the case-marking was not particularly acoustically salient, and participants may have misheard the accusative noun phrases as their nominative counterparts because of this nominative-first bias.[7] In constraint-satisfaction terms, the acoustic information may have only weakly constrained the interpretation of the acoustic signal as an accusative, whereas the bias to interpret a clause-initial noun phrase as nominative may have had an influence in the opposing direction.

In summary, we interpret Experiment 1 as providing evidence for the combination of real-world knowledge with both the morphosyntactic marking of the preverbal argument and the semantic constraints of the verb to enable the early anticipation of the postverbal referring expression.

EXPERIMENT 2

Experiment 2 used an experimental method and items similar to those of Experiment 1, but in a different language.

Participants

Thirty-two subjects from the University of York student community took part in this study. They were either paid or given part course credit for their participation. All were native speakers of British English and either had uncorrected vision or wore soft contact lenses or eyeglasses.

[6] Scheepers *et al.* (in prep.) report an eyetracking experiment using such verb manipulations.
[7] Scheepers *et al.* (in prep.) address this issue more closely in additional experiments.

Materials

The 16 experimental pictures used in Experiment 1 (e.g., Fig. 1) were used here. Each picture was paired with two sentential conditions: active—"The hare will eat the cabbage"(2a), or passive—"The hare will be eaten by the fox"(2b). The English sentences were produced mostly by literally translating the original German stimuli. However, a very small number of nouns and verbs needed some modification from the literal translation due to slight differences in the semantics of the words in the two languages. The postverbal adverbs in the German sentences (e.g., "gleich") were eliminated, as English does not easily permit a postverbal adverb in transitive constructions. The 16 sets of experimental items were divided into two lists in the same way as in Experiment 1. A further 16 items were added as fillers that served as experimental items for an unrelated experiment. Half of the fillers were active future-tensed sentences (e.g., "The cat will drink . . .") and the other half were active perfective-tensed sentences (e.g., "The cat has drunk . . ."). Each filler picture contained five objects in a semirealistic scene. The sentences were recorded and edited in the same way as in Experiment 1, except that they were recorded by a male native speaker of British English (GTMA). The experimental materials can be viewed at www.york.ac.uk/res/prg/jpr02.

Procedure

The procedure was predominantly the same as in Experiment 1, except for the following details. The display was 17″ (instead of 21″), and the resolution of the visual stimuli was 640 × 480 pixels (instead of 800 × 600). There was no comprehension task for participants to perform. Each participant's right eye was tracked. The auditory sentences were presented 1000 ms after the onset of the presentation of the pictures. Each picture was on the screen for a total of 6000 ms.

Results

The only crucial region in this experiment was between the onset of the verb and the onset of the second noun phrase, as the items before the verb onset were identical between conditions. Figure 3 shows the "verb" region used in the analyses.[8,9] Table II shows the percentages of trials where there

[8] The theoretically latest offset of the critical verb region would be the onset of the noun in the second noun phrase ("cabbage"/"fox"). However, as in Experiment 1 (and to be coherent with Experiment 1), we shall adopt the conservative definition of the region, not including eye movements that occurred during the determiner.

[9] In addition, we performed analyses on eye movements that occurred between the onset of NP1 and the onset of the verb. No reliable effect was obtained in this region.

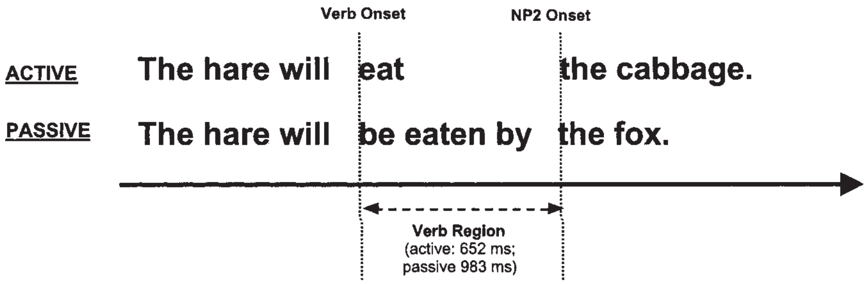

Fig. 3. The "verb" region used for the data analyses in Experiment 2. The numbers in the brackets show the mean duration of the region in each condition in milliseconds. The relative sizes of the space between the words in Fig. 3 do not correspond to the actual relative lengths of the pauses in the auditory materials.

was at least one fixation onto the "cabbage" objects or the "fox" objects in each sentential condition during the Verb region.

An ANOVA of the data in the Verb region found main effects of both Voice (passive > active; $F1(1,31) = 60.36$, $p < .001$; $F2(1,15) = 27.18$, $p < .001$) and Object (cabbage > fox; $F1(1,31) = 35.65$, $p < .001$; $F2(1,15) = 8.64$, $p < .02$). Most crucially, the interaction between Voice and Object was significant ($F1(1,31) = 5.46$, $p = < .05$; $F2(1,15) = 10.90$, $p < .005$). There were more looks to the "*appropriate*" objects (the cabbage in the active condition, the fox in the passive condition) than the "*inappropriate*" objects (the fox in the active condition, the cabbage in the passive condition) during this region. Planned comparisons revealed that there were significantly more looks toward the "fox" objects in the passive condition than in the active condition ($F1(1,31) = 24.28$, $p < .001$; $F2(1,15) = 41.47$, $p < .001$). However, there was no effect of voice on looks toward the "cabbage" objects ($F1(1,31) = 2.64$, $p > .10$; $F2(1,15) = 3.13$, $p < .10$). We conducted a further set of analyses to compensate for the fact that the verb region was 331 ms longer in the passive condition than in the active condition. For the active condition, we took eye movements that occurred between the onset of the verb ("eat") and

Table II. Percent of Trials with a Fixation onto the "Cabbage" Objects and the "Fox" Objects Obtained in Each Condition for the Verb Region in Experiment 2. The "Cabbage" Objects were the "Appropriate" Objects in the active Condition, and the "Fox" Ones for the Passive Condition

Region	Condition	"Cabbage" objects	"Fox" objects
Verb region	Active	31	11
	Passive	40	35

the onset of the determiner of the second noun phrase (a duration of 652 ms). In the passive condition, we took eye movements that occurred between the onset of the verb region ("be") and 652 ms beyond. The offset of this "Modified Verb" region for the passive condition typically fell somewhere in the past participle verb itself ("eaten"). Table III gives the percentages of trials with a fixation for this new region.

An 2 $\times$ 2 ANOVA yielded a main effect of Object (cabbage $>$ fox; $F1(1,31) = 21.63$, $p < .001$; $F2(1,15) = 15.57$, $p < .005$), but not of Voice ($F1$, $F2 < 1$). The interaction between the two was significant ($F1(1,31) = 4.15$, $p = .05$; $F2(1,15) = 10.269$, $p < .01$). Planned comparisons revealed a difference in looks to the "fox" objects as a function of Voice that was significant by items and approached significance by subjects ($F1(1,31) = 3.70$, $p < .10$; $F2(1,15) = 9.71$, $p < .01$). There was no difference in looks to the "cabbage" objects ($F1 < 1$; $F2(1,15) = 2.00$, $p > .10$). Thus, these new analyses confirm the effect of Voice on looks to the "fox" objects in the absence of a length confound across the conditions.

Discussion

Experiment 2 produced a similar pattern of data to Experiment 1: People looked at the "fox" objects more often when they were appropriate (the passive condition) than when they were inappropriate (the active condition), and they did so before the onset of "the fox." This suggests that prediction of the postverbal noun phrase was accomplished using a combination of voice information, the verbs' semantic constraints, and real-world knowledge. It follows that both syntactic and semantic constraints combine to enable the processor to access real-world knowledge for the purposes of predicting what will come next.

Experiment 2 also replicated the asymmetry in the data as seen in Experiment 1, with no more looks to the "cabbage" objects in the active condition than in the passive condition. As discussed in connection with Experiment 1, there are a number of reasons why there may have been, in effect, too many fixations on the "cabbage" objects in the accusative condi-

Table III. Percent of Trials with a Fixation onto the "Cabbage" Objects and the "Fox" Objects Obtained in Each Condition for the Modified Verb Region in Experiment 2

Region	Condition	"Cabbage" objects	"Fox" objects
Modified verb region	Active	30	13
	Passive	27	20

tions. Visual properties of the scenes might have attracted disproportionate numbers of looks toward the "cabbage" objects. On the other hand, the nominative-first preference in German is a straightforward case of Bever's "main clause strategy" in which the first noun phrase in a clause is interpreted as an Agent (e.g., Bever, 1970; Townsend & Bever, 2001). Given such a strategy, we would expect "The hare . . ." in both conditions to be interpreted as Agent, with the consequence that even in the passive condition, one would see some residual effect during the verb region of increased looks toward an appropriate Theme given the hare as Agent. Overall, we believe the data, and specifically those concerning looks toward the "fox" objects, to show evidence of anticipatory processing.

GENERAL DISCUSSION

The main motivation underlying Experiments 1 and 2 was to see the effect of case-marking information (or the lack of it) on the time-course of prediction in two languages with different role/function designation systems. In the German experiment, the first noun phrase was explicitly case marked, which indicated the thematic role to be assigned to the head noun. In contrast, in English, information about thematic roles does not become available until the verb is encountered (notwithstanding an Agent-first strategy). Interestingly, significant anticipatory eye movements were obtained at the verb in the English study, but only in the *following* region in the German study. We note, however, that the German verb region was on average shorter than its English counterpart (516 ms vs. 739 ms). Perhaps the discrepancy in the time-course of anticipatory processing is simply a reflection of a longer region in which to move the eyes in the English case relative to the German case. To explore this possibility, we conducted further analyses of the English data in which we eliminated eye movements that occurred within 223 ms of the onset of the postverbal noun phrase (its determiner)—thereby making the analysis region exactly the same duration, on average, as the German verb region. The additional analyses maintained the same cross-linguistic difference: This new analysis of the English data still gave the same pattern of anticipatory eye movements as before.

Thus, the question remains: Why did we get this rather counterintuitive cross-linguistic difference? We believe that minor variations in the experimental procedures are not a plausible candidate for explaining this difference (unless the additional 200 ms viewing time that the German participants were given somehow interacted with the auditory stimuli to delay processing). Rather, we speculate that integrating different types of information across different constituents requires more processing time than integrating different types of information from a single constituent. Note that the crucial syntactic information in the German study was extracted from the first noun phrase and the crucial semantic information from the verb; in the English study both types of information were extracted at the verb. Thus, in German, by the time the processor reaches the second information source (the verb) the first piece of information regarding case information has already appeared and perhaps decayed to a degree (cf. Altmann *et al.*, 1998; Stevenson, 1994), and any such decay may engender later anticipatory eye movements than in conditions where no such decay occurs. Naturally, our study does not clarify whether this is a simple distance effect (distance between the information sources that together enable the predictive process) or whether it is an effect that is dependent less on distance and more on integration across constituents that have different syntactic types. Nonetheless, we believe that our cross-linguistic study suggests that integrating information across multiple sources during sentence processing is not without cost, even if the source of that cost is yet to be identified.

REFERENCES

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73,* 247–264.

Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191–238.

Altmann, G. T. M., van Nice, K., Garnham, A., & Henstra, J. A. (1998). Late closure in context. *Journal of Memory and Language, 38,* 459–484.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419–439.

Bader, M., & Lasser, I. (1994). German verb-final clauses and sentence processing: Evidence for immediate attachment. In C., Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on Sentence Processing* (pp. 225–242). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language,* (pp. 273–312). New York, NY: Wiley.

Chambers, C. G., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real time language comprehension. *Journal of Memory and Language, 47,* 30–49.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84–107.

Duebel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827–1837.

Eberhard, K., Spivey-Knowlton, S., Sedivy, J., & Tanenhaus, M. (1995). Eye movements as a window into real-time spoken language processing in natural contexts. *Journal of Psycholinguistic Research, 24,* 409–436.

Frazier, L. (1979). *On Comprehending Sentences: Syntactic Parsing Strategies.* Bloomington, IN: IULC.

Hemforth, B., & Konieczny, L. (Eds.) (2000). *German Sentence Processing* (Studies in Theoretical Psycholinguistics, Vol. 24). Dodrecht: Kluwer Academic Publishers.

Hoffman, J. E., & Surbramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57,* 787–795.

Kako, E., & Trueswell, J. C. (2000). Verb meanings, object affordances, and the incremental restriction of reference. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 256–261). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kamide, Y., Altmann, G. T. M., & Haywood, S. (in press). Prediction and thematic information in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language.*

Kamide, Y., & Mitchell, D. C. (1999). Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes, 14,* 631–662.

Koh, S. (1997). The resolution of the dative NP ambiguity in Korean. *Journal of Psycholinguistic Research, 26,* 265–273.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35,* 1897–1916.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.

Marlsen-Wilson, W. D. (1975). Speech perception as an interactive parallel process. *Science, 189,* 226–228.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance, XII* (pp. 601–618). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pollard, C., & Sag, I. A. (1987). *Information-Based Syntax and Semantics.* CSLS lecture notes (No. 13). Stanford: Centre for the Study of Language and Information, 1.

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar.* Chicago, London: University of Chicago Press and CSLI Publications.

Pritchett, B. L. (1991). Head position and parsing ambiguity. *Journal of Psycholinguistic Research, 20,* 251–270.

Pritchett, B. L. (1992). *Grammatical Competence and Parsing Performance.* Chicago: University of Chicago Press.

Rayner, K., McConkie, G. W., & Ehrlich, S. F. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 529–544.

Remington, R. W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 726–744.

Scheepers, C., Kamide, Y., & Altmann, G. T. M. (in prep.). The integration of verb restrictions and case information in projecting upcoming arguments in German: Evidence from anticipatory eye movements.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71,* 109–148.

Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research, 23,* 295–322.

Sussman, R. S., & Sevidy, J. C. (2001, September). Argument structure and the time course of gap-filling processes. Paper presented at the AMLaP-2000 conference, Leiden, Netherlands.

Tanenhaus, M. K., Garnsey, S., & Boland, J. (1990). Combinatory lexical information and language comprehension. In Altmann, G. T. M. (Ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives* (pp. 271–408). Cambridge, MA: MIT Press.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. G. (2001). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29,* 557–580.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Townsend, D. J., & Bever, T. G. (2001). *Sentence Comprehension: The Integration of Habits and Rules.* Cambridge, MA: MIT Press.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73,* 89–134.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preferences from garden-paths. *Journal of Experimental Psychology: Learning, Memory & Cognition, 19,* 528–553.