Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 49 (2003) 133–156
Journal of Memory and Language
www.elsevier.com/locate/jml

The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements

Yuki Kamide,* Gerry T.M. Altmann, and Sarah L. Haywood

Department of Psychology, University of York, UK

Received 18 October 2001; revision received 24 January 2003

Abstract

Three eye-tracking experiments using the 'visual-world' paradigm are described that explore the basis by which thematic dependencies can be evaluated in advance of linguistic input that unambiguously signals those dependencies. Following Altmann and Kamide (1999), who found that selectional information conveyed by a verb can be used to anticipate an upcoming Theme, we attempt to draw here a more precise picture of the basis for such anticipatory processing. Our data from two studies in English and one in Japanese suggest that (a) verb-based information is not limited to anticipating the immediately following (grammatical) object, but can also anticipate later occurring objects (e.g., Goals), (b) in combination with information conveyed by the verb, a pre-verbal argument (Agent) can constrain the anticipation of a subsequent Theme, and (c) in a head-final construction such as that typically found in Japanese, both syntactic and semantic constraints extracted from pre-verbal arguments can enable the anticipation, in effect, of a further forthcoming argument in the absence of their head (the verb). We suggest that such processing is the hallmark of an incremental processor that is able to draw on different sources of information (some non-linguistic) at the earliest possible opportunity to establish the fullest possible interpretation of the input at each moment in time.

An extensive body of work has been carried out to address the extent to which the human sentence processor processes its inputs incrementally; numerous recent sentence processing models have supported incrementality at some level. It is commonly held, for instance, that the incoming inputs are processed without delay on a word-by-word basis, that relevant constraints are applied in parallel to the analysis of the input as it unfolds, and that all relevant analyses of the unfolding input are specified to some degree (e.g., Altmann & Steedman, 1988; MacDonald, Pearlmutter, & Seidenberg, 1994). The present paper follows Altmann and Kamide (1999) in exploring the extent to which the word-by-word analysis of a sentence can lead to the assignment of thematic roles before the point in the linguistic input at which that assignment is unambiguously signaled. Specifically, the studies described below establish that contextual information and combinatory information—information combined from different lexical sources—can be applied as soon as those lexical sources are encountered (that is, at the theoretically earliest moment in time), and that the application of these distinct sources of information enables the processor to anticipate, again at the theoretically earliest moments, what will be referred to next.

Early studies on incremental processing explored either the immediacy with which contextual information might influence the earliest stages of syntactic ambiguity

* Corresponding author. Present address: Department of Psychology, University of Manchester, Oxford Road, Manchester M13 9PL, UK. Fax: +44-0-161-27-52685.
E-mail address: yuki.kamide@man.ac.uk (Y. Kamide).

0749-596X/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0749-596X(03)00023-8

resolution (e.g., Altmann & Steedman, 1988; Altmann, Garnham, & Henstra, 1994; Mitchell, Corley, & Garnham, 1992; Tyler & Marslen-Wilson, 1977) or the extent to which lexically derived information might 'drive' the parsing process. For example, Tanenhaus, Carlson, & Trueswell (1989) assumed that the recognition of a verb entails the specification of its meaning, of the thematic roles associated with that meaning, and of the syntactic constituents that manifest those roles in the syntactic structure of the sentence. They suggested that this information could enable the processor to 'project an expectation about what kinds of linguistic elements occur after the verb' (p. 214), and that a provisional meaning could be assigned to the sentence in the absence of those linguistic elements—the particular identities that would be assigned to those elements would be left unspecified until such time as they were encountered, although the roles that they play would be specified in the provisional meaning assigned to the sentence. They suggested also that 'if a likely entity has already been introduced in the discourse or even the same sentence, it may be provisionally assumed to be playing one of [the verb's thematic] roles until further evidence shows otherwise.' (p. 214).

In respect of the latter claim, the authors and their colleagues demonstrated that long-distance filler-gap dependencies appear to be resolved on the basis of precisely such assumptions. Thus, Boland, Tanenhaus, Garnsey, and Carlson (1995) investigated the processing of wh-questions such as *Which preschool nursery did Hank deliver the machine guns to _ last week?* and argued that the processor anticipates at the verb *deliver* that the preschool nursery will serve its Goal role; they found more 'stops making sense' judgments at *machine guns* in this sentence than in the sentence *Which* military base *did Hank deliver the machine guns to _ last week?*; they accounted for this pattern by assuming that in the first case the processor assigns the Goal role to the preschool nursery at the verb, and that this causes the subsequent *machine guns* to be deemed an implausible Theme given the nursery as Goal (in contrast to the second case where they are a plausible Theme given the military base as Goal).

More recently, Altmann (1999) showed equivalent effects in the absence of any obligatory wh filler-gap dependencies. He contrasted the following two cases: *Hank parked his van outside the preschool nursery. He delivered some machine guns to the military base next door.* and *Hank parked his van outside the local military base. He delivered some machine guns to the military base and left.* Like Boland et al. (1995), he observed more 'stops making sense' judgments (and longer reading times) at *machine guns* in the case where the preschool nursery had been mentioned than in the case where a military base had been mentioned. Unlike the Boland et al. (1995) study, there is no syntactic dependency in these cases on which basis the assignment of the Goal role to the nursery is licensed. Nonetheless, it appeared that the processor again anticipated the assignment of the Goal role, presumably at the verb, and on the basis that a discourse entity existed that could plausibly play the Theme role for the verb *deliver*.' Taken together, the results support Tanenhaus and colleagues' original contention that lexical information about the verb coupled with discourse-based information about available entities, can enable the processor to anticipate thematic role assignments; that is, to make thematic role assignments at the verb or soon after that correspond to syntactic dependencies between that verb and a post-verbal constituent that has not yet been encountered.

We shall refer to this anticipation of thematic role assignment as a form of 'prediction' (insofar as knowledge of the grammar would enable the processor, in principle, to predict an upcoming constituent that should receive the assigned role); we shall return in General Discussion to discussion of the content of the representations that constitute these predictions, and specifically, whether any part of these representations corresponds to the projection, forward in time, of linguistic structure (syntactic or semantic). For now, however, we use the terms 'anticipation,' 'prediction' and 'predictive processing' interchangeably, and in the context of thematic role assignment, to refer to the process by which thematic dependencies can be established prior to the point at which those dependencies are unambiguously signaled within the linguistic input.

The investigation of predictive processes during sentence processing can shed light on the different types of information that can guide parsing. Numerous syntactic theories maintain that a lexical head contains crucial information about its permissible arguments (e.g., Chomsky, 1965, 1981, 1995; Pollard & Sag, 1994), and most discussion has centered around the verb within these frameworks. Thus, it is important to test empirically how such information is used as the head and its arguments are encountered within the unfolding sentence. Many studies have demonstrated the use of information stored with verbs during the processing of head-initial languages, where internal arguments of a verb are encountered after the verb (e.g., Altmann, 1999; Altmann & Kamide, 1999; Boland, Tanenhaus, & Garnsey, 1990; Boland et al., 1995; Trueswell, Tanenhaus, & Kello, 1993; McElree & Griffith, 1995; but see also Mitchell, 1987). One important question is whether the constraints that a verb places on its potential objects (e.g. whether such objects are grammatically licensed, whether they must conform to certain semantic restrictions, and so on) come into effect as soon as the verb is encountered (and consequently, whether the potential objects, and their properties, are anticipated as soon as the verb appears) or whether the application of such constraints is delayed until the object is encountered

(in which case some sort of checking procedure might evaluate the goodness of fit between that grammatical object and the semantic specification of the verb). Hence, investigations of predictive processing can help clarify at what point during the processing of a sentence information extracted at the verb can begin to guide the parsing processes, and consequently what types of information are used when.

Previously (in Altmann & Kamide, 1999), we have explored whether the anticipation of upcoming material can be made on the basis of *selectional* restrictions—semantic requirements imposed upon a noun to be qualified as an argument of the verb (Chomsky, 1965; Jackendoff, 1977). For example, the verb *drink* requires its direct object (Theme) to be something drinkable, often liquid.[1] Within some theoretical frameworks, such constraints would not naturally play a guiding role in predicting semantic properties of verb's subsequent arguments, even if the arguments themselves might be predicted, because semantic constraints of this kind are not stored as part of the information associated with the head of a phrase (e.g., HPSG: Pollard & Sag, 1994; but also see Androutsopoulos & Dale, 2000, for an implementation of selectional restrictions into HPSG). In other theoretical frameworks, semantic constraints on a verb's arguments would not be applied during the prediction of those arguments because the access of such semantic information is delayed until after syntactic information is used (e.g., Rayner, Carlson, & Frazier, 1983). Nonetheless, these different frameworks are not incompatible with accounts that suggest that a verb's arguments are projected at the verb, and that by the time those arguments are encountered in the linguistic input, their projections have been evaluated semantically. The purpose of our own research (cf. Altmann & Kamide, 1999) is to establish whether indeed such evaluation takes place prior to encountering the verb's arguments. Again, we leave until later a discussion of whether the projection of syntactic structure necessarily accompanies the anticipation of a verb's arguments; we are more concerned at present with the nature of the information that informs the anticipatory process.

Altmann and Kamide (1999) explored this issue in the context of selectional restrictions associated with a verb in simple subject–verb–object sentence frames (although what we refer to here as 'selectional restrictions' may be no more than semantic constraints reflecting more general experiential knowledge—cf. McRae, Ferretti, & Amyote, 1997). Specifically, Altmann and Kamide explored whether the processor would anticipate, at the verb, a subsequent object with semantic properties constrained by the selectional restrictions. They used the visual-world paradigm initiated by Cooper (1974) and further developed by Tanenhaus and colleagues (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). In their first experiment, Altmann and Kamide presented an auditory sentence such as *The boy will eat the cake.* while participants viewed a scene containing a boy, a cake, a toy train, a toy car, and a ball. In this scene, the cake was the only object that satisfies the selectional restrictions of the verb (something edible: the selective condition). A second condition, the non-selective condition, used sentences such as *The boy will move the cake.* presented with the same scene. In this condition, the selectional restrictions no longer exclusively selected for the cake: the other objects in the scene—the toy train, the toy car, and the ball—were all moveable, and thus satisfied the semantic restrictions of the verb. During the experiment, participants were asked to judge whether or not the sentence could apply to the entities represented in the scene. Altmann and Kamide found that the probability of launching a saccadic eye movement to the target object was significantly higher in the selective condition than in the non-selective condition, and that this difference manifested itself even before the actual noun ('cake') was encountered; the difference reached statistical significance well before the onset of the expression *the cake*. These data were interpreted as suggesting that the processor anticipates at the verb a forthcoming post-verbal argument, applies to this argument the semantic constraints afforded by the verb's selectional restrictions, and evaluates the result against, in this case, the visual context. In effect, therefore, selectional restrictions at the verb can be used to restrict the domain of subsequent reference, even before the post-verbal referring expression is in fact encountered within the linguistic input.

The present paper adopts the paradigm and technique used in Altmann and Kamide (1999) in order to further explore thematic role assignments in context: we use the same visual world paradigm and monitor for *anticipatory eye movements* towards certain target objects, and we shall take such eye movements as evidence for predictive processing: Given the concurrence of the visual scene and the auditory presentation of the linguistic stimulus, it is possible to monitor movements towards objects *before* the referring expressions for those objects have been encountered in the linguistic input (hence the term 'anticipatory' eye movements). If the eyes move to one object more than to another as a function of the linguistic input, we can infer the time-course with which that input was interpreted from the temporal relationship between the onset of the eye

[1] Throughout the present paper, we shall use the term 'Theme' for the direct object, 'Goal' for the indirect object, and 'Agent' for the grammatical subject of the sentence (although grammatical and thematic roles can be distinguished in some other cases, they coincide in our sentences). In addition, where a referring expression that would receive the Theme role refers to a particular object in an accompanying visual scene, we shall refer to that object as the Theme object.

movement and the onset of the linguistic input that determined the occurrence of that movement. And although most normal language processing neither requires nor intends overt shifts in visual attention, it is undeniably the case that language can and is used to cause such shifts.

We describe below three experiments that extend previous findings on thematic role assignment in context. Experiment 1 is a visual analogue of the Altmann (1999) study referred to earlier. This version uses the contrast between *The woman will spread the butter on the bread* and *The woman will slide the butter to the man* to explore whether information associated with a 3-place verb can be used to anticipate properties of the second post-verbal argument (typically Goal—the bread or the man) prior to its onset. In some respects, this is also a simple extension of the Altmann and Kamide (1999) study, which explored whether information associated with a monotransitive verb can be used to anticipate properties of the immediately post-verbal argument (typically Theme or Patient) prior to its onset. However, whereas in that study anticipatory eye movements were found during or just after the verb (and so during a portion of the sentence that was not explicitly referential), in Experiment 1 we monitor during the first post-verbal argument (*butter*) for anticipatory eye movements towards the object referred to by the second post-verbal argument (the bread or the man)—will anticipatory eye movements reflecting anticipation of a later referring expression be manifest during an earlier referring expression? Experiment 2 uses monotransitive verbs to explore whether only information associated with the main verb can drive the anticipation of what will plausibly be referred to next, or whether the *combination* of this information with the prior grammatical subject can drive such predictions. Anticipatory eye movements in Experiment 1 towards, for example, the man in *The woman will slide the butter*... could be due simply to the fact that a verb like *slide* might more often be accompanied by an animate Goal than an inanimate Goal (or Location). Thus, anticipatory looks to the man might not necessarily be contingent on the assignment of the Theme role to any particular entity. The purpose of Experiment 2 is to establish whether such contingent effects *are* possible. To this end, Experiment 2 keeps the verb the same but manipulates the subject (cf. *The man* will ride the motorbike. vs. *The girl* will ride the carousel.'). At issue here is whether we shall observe anticipatory eye movements at the verb to just those entities in the scene that are plausibly ridden, or to just those entities in the scene that are plausibly ridden by whoever is the subject of the verb. Finally, whereas the data from Experiments 1 and 2 will be explained in terms of sensitivity to real-world contingencies, Experiment 3 will explore whether morphosyntactic information, as opposed to information pertaining only to real-world contingencies, can also inform predictive processes during sentence processing. This final study employs the verb-final language Japanese to explore whether the first two arguments of a 3-place verb can predict the third (cf. *the waitress will bring the burger to the customer*, which glosses as *waitress customer burger bring*). This study also explores whether aspects of the semantics of the verb are predicted on the basis of those first two arguments.

Experiment 1

In this experiment, we consider the processing of 3-place verbs, as in

(1) The woman will spread the butter on the bread.
(2) The woman will slide the butter to the man.

The verbs *spread* and *slide* subcategorize for up to two post-verbal arguments. At issue is whether the processor anticipates, either at the verb or during the first post-verbal argument (the Theme), information pertaining to the second post-verbal argument (the Goal). This study is analogous to the Altmann (1999) experiment, except that, in the present experiment, the discourse entity that we hypothesize will be anticipated at the verb to receive the Goal role is introduced not in a prior linguistic context, as in that earlier study, but in a concurrent visual context: In the context of a visual scene portraying a woman, a man, butter, and bread, we would expect to observe anticipatory eye movements towards the bread during *the butter* in (1) above (the bread is a plausible Goal given the butter as the Theme of *spread*), but towards the man during *the butter* in (2), where the man is a plausible Goal given the bread as the Theme of *pass*. Whether we might also observe anticipatory eye movements towards the bread or the man during the verbs themselves is unclear, although we shall return to this point in the discussion of the results. Although Altmann and Kamide (1999) demonstrated that the processor can predict the Theme in a monotransitive construction (equivalent to anticipating the butter after *spread*), there are reasons to suppose that Goals may not be predicted in 3-place constructions to the same extent. First, Goals tend to appear later than Themes in English, and within an interactive activation framework, constraints extracted at the verb might decay over time such that they are no longer effective predictors by the time the second post-verbal argument is due (cf. Altmann, Garnham, van Nice, & Henstra, 1998; MacDonald et al., 1994). Second, the majority of 3-place verbs in English only optionally take two post-verbal arguments; there are very few obligatorily 3-place verbs in English. Consequently, overt (or explicit) Goals are considerably rarer than Themes (there are biases, with certain potentially 3-place verbs 'preferring' a Theme and a Goal, and others preferring just a Theme; Kamide

& Altmann, in preparation). On this basis, the processor may be less likely to predict a Goal after a Theme.

Experiment 1 also raises the question of whether anticipatory eye movements are restricted to moments in time when the linguistic input does not explicitly refer to any object in the visual scene. For instance, in Altmann and Kamide (1999), significant anticipatory eye movements were obtained before the onset of the post-verbal referring expression *the cake*. Perhaps these anticipatory eye movements were observed because, in effect, there was nothing else that the language was explicitly referring to at this time (during the verb and immediately beyond)—eye movements might generally reflect only what is being referred to in the here-and-now, and might only reflect what might be referred to next when there is nothing being referred to in the here-and-now. In Experiment 1, we monitored for anticipatory eye movements during the post-verbal referring expression *the butter*.

Additionally, Experiment 1 aims to address another issue, relating to the methodology. Altmann and Kamide's (1999) first experiment required participants to judge whether the event described in the sentence could occur given the objects portrayed in the scene. In half the trials (filler trials), therefore, an object was mentioned that did not appear in the scene (hence allowing a 'no' response). Their second experiment used the same material set but a different experimental task and instructions. In this second experiment, participants were not asked to perform the sentence-picture judgement task (they simply had to look at the scene and listen to the sentence), and although the effects became significant somewhat later in the second experiment than in the first (but still significant by the onset of *the cake*), the pattern of eye movements replicated the earlier result, suggesting the robustness of the prediction process as evidenced in this paradigm. However, the presence of 'mismatching' trials may have led participants to attempt to anticipate whether the sentences would match or not, and it is thus possible that the participants did implicitly validate each sentence against its corresponding picture. It is conceivable, therefore, that the mismatched trials caused participants to adopt a strategy that was effectively the same as in the overt judgment task used in Altmann and Kamide's (1999) Experiment 1. If that is the case, participants may have converged on anticipatory processing strategies that did not reflect normal processing, but rather reflected instead the exigencies of the (implied) judgment task—in order to perform the judgment quickly, participants may have recruited information (verb selectional restrictions) that are not usually recruited during normal sentence processing. In Experiment 1, we used only matching picture/sentence pairings, and told participants that we were 'interested in what happens when people look at these pictures while listening to sentences that describe something that might happen in the picture'.

Experiment 1 takes advantage of the fact that many English 3-place verbs place semantic constraints on their Goals. Thus, in (1) and (2) reproduced below, the bread is a more plausible Goal of the spreading than a man would be, and the man is a more plausible Goal for the sliding. The two sentential conditions used in the experiment corresponded to conditions in which an inanimate Goal was preferred or an animate Goal.

(1) *Inanimate Goal*. The woman will spread the butter on the bread.
(2) *Animate Goal*. The woman will slide the butter to the man.

Fig. 1, panel A shows the visual scene that was presented concurrently with either sentence (1) or (2). Fig. 1, panel B indicates the critical regions for data analyses ('Region 1' and 'Region 2' as we shall discuss below).

We hypothesized that 'appropriate' Goals should be fixated on more often before the onset of the preposition (*on* or *to*) than the 'inappropriate' Goals (appropriate: bread in (1) or man in (2); inappropriate: bread in (2) or man in (1)). If such an eye movement pattern is obtained, this would show first that the semantic restrictions imposed by the verb remain effective enough to predict the second post-verbal argument (the Goal). Second, if the pattern of anticipatory eye movements is found specifically during the lifetime of the head of the Theme (*butter*), this would rule out the possibility that predictive eye movements only occur during linguistic material that is not explicitly referential. Third, such a result would demonstrate that anticipatory eye movements can be obtained even if there is no explicit task for the participant to perform other than to look and listen (and when all sentence–picture pairings match).

Method

Subjects

Sixty-four subjects from the University of York student community took part in this study. They participated either for course credit or for £2.00. All were native speakers of English and either had uncorrected vision or wore soft contact lenses or spectacles.

Stimuli

There were 18 experimental pictures (e.g., Fig. 1) each paired with two sentential conditions (e.g. (1) and (2) above). The visual scenes were created using commercially available ClipArt packages, and were constructed using a 16-colour palette. They were presented on a 17 in. viewing monitor at a resolution of 640×480 pixels. Each scene contained five objects: Agent, Theme, Inanimate Goal, Animate Goal, and a Distractor. A further 24 sentences were added as fillers. They were all monotransitive constructions, and all consisted of picture–sentence pairings that matched. The materials were arranged in a fixed–random order so that no successive

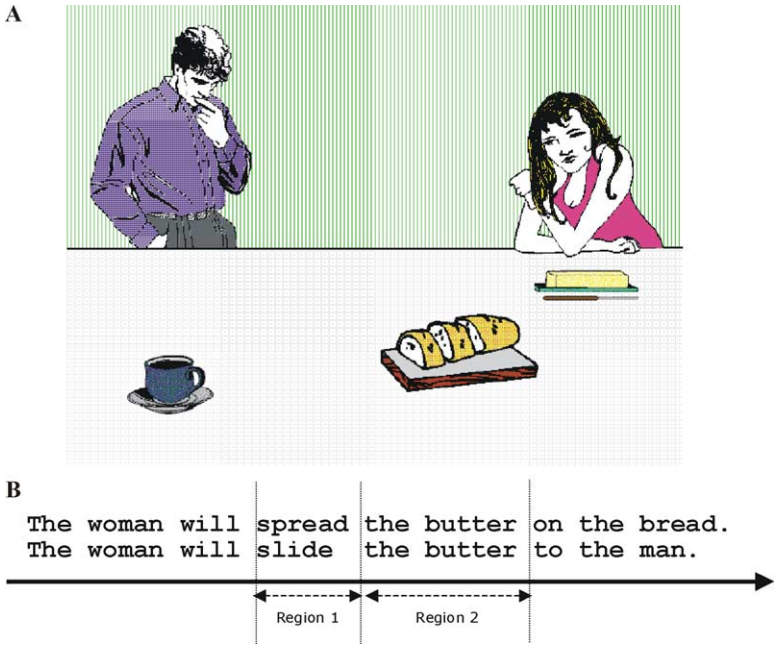

Fig. 1. An example visual (panel A) and auditory (panel B) stimuli from Experiment 1. The regions relevant to the analyses are indicated as Region 1 and Region 2. For the auditory stimuli, the relative sizes of the space between the words do not correspond to the actual relative lengths of the pauses in the sentences.

items belonged to the same condition. Two lists of stimuli were created containing each of the 18 experimental pictures but just one version of each sentence pair—nine of the 18 sentences were drawn from the animate condition, and nine from the inanimate condition.

The sentences were recorded by a male native speaker of British English (GTMA), and sampled at 44,100 Hz. The sound files were presented to participants via a mono channel split to 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. The experimental sentences for Experiment 1 are given in Appendix A.

Procedure

Participants were seated in front of a 17 in. display with their eyes between 20 and 25 in. from the display. They wore an SMI EyeLink head-mounted eye-tracker, sampling at 250 Hz from the right eye (viewing was binocular). Participants were given the following instruction: 'In this experiment we shall ask you to look at some pictures on a computer screen. Each picture will be accompanied by a short sentence spoken over the loudspeakers. The picture might show, for example, a person in a room with a fireplace and various objects. The sentence might be like 'the person will light the fire'. We are interested in what happens when people look at these pictures while listening to sentences that describe something that might happen in the picture.' There were two practice trials before the main experimental block. Between each trial, participants were shown a single centrally located dot on the screen that they were asked to fixate prior to a fixation cross appearing in this position (this procedure allowed mini-recalibration of the eye-tracker). Participants would then press a response button for the next presentation. The onset of the visual stimulus coincided with the onset of the spoken stimulus (both were stored and played from disk). The trial was automatically terminated after 6 s, which typically left 2–3 s after the end of the sentence. The termination of trials was pre-set and controlled by the experimental program, so that participants could not terminate trials by themselves. After every fourth trial, the eye-tracker was recalibrated using a 9-point fixation stimulus. The EyeLink software automatically validates calibrations and the experimenter could, if required, repeat the calibration process if validation was poor. Calibration took approximately 20 s. The entire experiment lasted approximately 25 min.

Results

We adopted the same procedure for analyzing the eye movement data generated by the EyeLink system as described in Altmann and Kamide (1999). We sought to determine when the participants' eyes first moved to the regions corresponding to the Appropriate or Inappropriate objects. Eye movements that landed more than

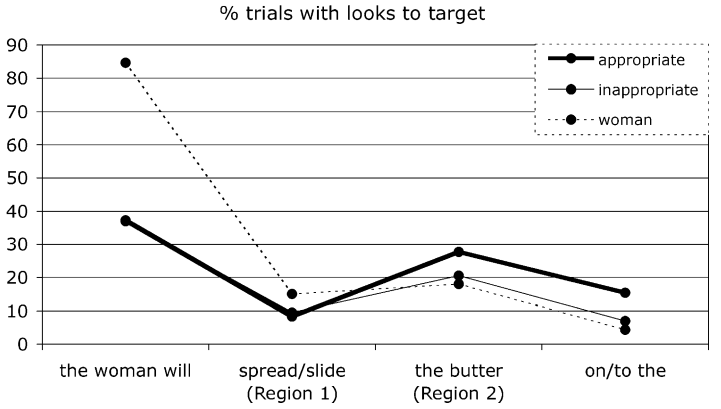

Fig. 2. Percentage of trials for each region of interest with looks to the appropriate target, the inappropriate target, and the Agent (the woman, in this example).

one or two pixels beyond the boundaries of the target object were not counted as fixations on that object. For the purposes of statistical analysis, we analyzed data in two different temporal regions as shown in Fig. 2. *Region 1*: the onset of the verb (*spread/slide*) to the onset of the Theme expression (*the butter*). *Region 2*: the onset of the Theme expression (*the butter*) to the onset of the preposition (*on/to*). The overall mean duration for each of these regions was: Region 1—350 ms; Region 2—882 ms.

The mean numbers of trials in which fixations onto both the Appropriate objects and the Inappropriate objects occurred were calculated for each region (only 'first' fixations within that region were used in the analyses, i.e., re-fixations in the same region were eliminated). Saccadic movements whose onset was prior to the region onset were eliminated from the analyses for that region. For example, any saccadic movements starting before *spread/slide* were not included in the data for Region 1. Objects in the scene pictures were labeled as in the following example, Appropriate objects: the bread for the Inanimate condition, the man for the Animate condition; Inappropriate objects: the bread for the Animate condition, the man for the Inanimate condition. While it might at first appear natural to compare, for example, anticipatory looks to the man during *slide the butter*... with looks to the bread, there are several reasons other than the experimental manipulation that may lead to one object in the scene attracting more looks than another. For example, relative position, color, contrast, and even spatial frequency may favor certain objects. And aside from the objects' different physical characteristics, their 'semantics' may also determine the pattern of looks—we have observed in a range of similar studies substantially more looks towards animate objects than towards inanimate ones. In this study, for example, participants looked at the man in the region between sentence onset and verb onset (*The woman will*...) on 53% of trials, but they looked at the bread on only 22% of trials. A final objection to comparing looks to different objects in the same scene/trial is that comparisons of such looks would violate the normal assumptions of statistical independence between conditions (looks to one object would necessarily be dependent on whether or not the participant looked at the other object). It is for these reasons that we analyze looks in terms of Appropriate and Inappropriate objects—this ensures that we analyze, in effect, looks towards the *same* objects under linguistically *different* conditions (and of course, this comparison is equivalent to an interaction between verb type and object, but it has the advantage of avoiding the pitfall of comparing looks to different objects). All the following statistical analyses were performed on the arcsine-transformed values of the data.

Fig. 3 shows the mean percentages of trials in which there was a saccadic eye movement towards the Appropriate objects or the Inappropriate objects in the different regions of the sentence (both the preceding and following regions are graphed also). In this and subsequent graphs of this kind, the percentages were calculated by summing the number of trials in which a saccadic eye movement was directed towards the targets; this number was calculated on a trial-by-trial basis taking into account, for that specific trial, where the region began (e.g., onset of the verb) and where it ended (e.g., onset of the following determiner).

In Region 1 (*spread/slide* _—we use '_' to indicate that the region extended to the onset of the following word), there was no difference in the proportion of trials with anticipatory looks to the appropriate or inappropriate objects (8.3 and 9.5%, respectively: $F1 < 1; F2(1,17) = 2.04,\ p > .1$). The difference in Region 2 (*butter_*) was, however, significant by both subjects and items (27.8 and 20.6% respectively, $F1(1,63) = 16.6,\ p < .0002;\ F2(1,17) = 11.0,\ p < .005$). The exact same pattern was observed when the regions of interest were changed to *spread/slide the_* and *butter_* (so incorporating the

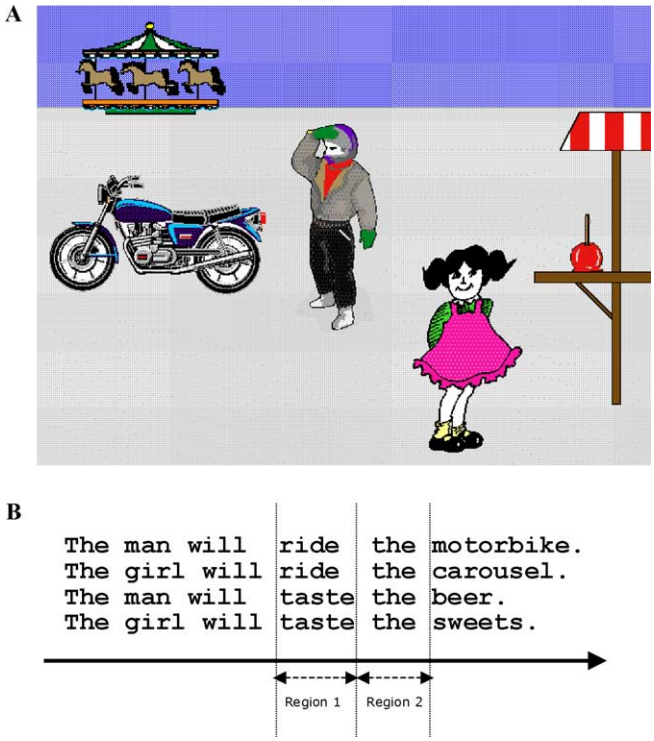

Fig. 3. An example visual (panel A) and auditory (panel B) stimuli from Experiment 2. The regions relevant to the analyses are indicated as Region 1 and Region 2. For the auditory stimuli, the relative sizes of the space between the words do not correspond to the actual relative lengths of the pauses in the sentences.

determiner into the verbal region). We also computed a 2×2 analysis of variance with factors Object (in effect, bread vs. man) and Verb (*spread* vs. *slide*). Overall, there were marginally significant effects of Verb ($F1(1,63) = 6.4,\ p < .02$; $F2(1,17) = 2.1,\ p > .1$) and of Object ($F1(1,63) = 4.7,\ p < .04$; $F2 < 1$) but a significant interaction between the two ($F1(1,63) = 16.5,\ p < .0001$; $F2(1,17) = 11.3,\ p < .004$). There were more trials containing at least one anticipatory look during *the butter_* to the man in the *slide* condition (24.3%) than to the man in the *spread* condition (16.8%)—($F1(1,63) = 20.8,\ p < .0001$; $F2(1,17) = 12.5,\ p < .003$). However the numerical difference in anticipatory looks to the bread (*slide*: 24.3%; *spread*: 27.4%) was not significant ($F1(1,63) = 1.4,\ p > .2$; $F2(1,17) = 1.5,\ p > .2$). We attribute this asymmetry to the fact that the bread (or its equivalent across trials) could be anticipated to be a potential Theme in the *slide* condition, meaning that the proportion of looks to the bread in this condition (24.3%) was inflated because looks had been directed towards it in anticipation that it would function as the Theme.

Notwithstanding this asymmetry, the data do demonstrate that for the animate objects at least, anticipatory looks did vary as a function of their appropriateness in the Goal role.

Discussion

The data demonstrate that during the immediately post-verbal referring expression (during '*the butter_*'—between the onset of the determiner and the onset of the subsequent preposition), there are more anticipatory eye movements towards the appropriate Goal object than towards the inappropriate Goal object. Given the 200 ms generally assumed necessary for programming and initiating an eye movement (e.g., Matin, Shao, & Boff, 1993; Saslow, 1967), these anticipatory eye movements must have been initiated during or very soon after the immediately post-verbal referring expression (the mean duration between the offset of *butter* and the onset of the preposition was 279 ms).

Our interpretation of these data is potentially complicated by the fact that more 3-place verbs in the condition with an animate Goal allowed an alternating constituent order than those in the condition with an inanimate Goal. For example, the verb *show* can appear in two forms: *The man will show the poster to the woman.* or *The man will show the woman the poster.*. But whereas *show* permits dative alternation, *tack* does not: *The man will tack the poster on the wall* vs. **The man will tack the wall the poster*. This alternating tendency seems to

coincide with the preference for an animate Goal. The relevance of this observation is that it casts doubt on whether the processor anticipated the Goal to be referred to in first or in second post-verbal position. Increased looks towards the man during *butter_*' may have occurred not because the processor anticipated that the man would be referred to next, but rather because the processor had anticipated that the man would be referred to in the position occupied by *butter*, with increased looks to the man in this position reflecting the processor's anticipation at *slide the* that the man would be referred to next. Our data do not rule out this possibility, and all we can conclude is that the data reflect verb-mediated anticipatory eye movements towards a plausible Goal—what is in doubt is whether the processor anticipated the Goal to occur only in the second post-verbal 'slot,' or whether it anticipated it to occur in the first post-verbal slot also. Within the probabilistic framework exemplified by MacDonald et al. (1994), it would have anticipated the Goal to occur in one or other post-verbal position as a function of the frequency, for that verb, of one or other argument structure.

In summary, the data demonstrate that the processor can anticipate Goal arguments. They demonstrate also that anticipatory eye movements to a particular object in the scene can be obtained even during reference to some other object in the scene, and that these anticipatory patterns can be obtained in a 'look and listen' task that does not require, nor imply, any meta-linguistic judgment.

While Experiment 1 was in some respects an analogue of the Altmann (1999) study, it was in other respects quite different: In that earlier study (*He delivered some machine guns...*), it was the assignment of a Goal role licensed by the verb (*deliver*) to an available discourse entity (the preschool nursery) that caused the post-verbal *some machine guns* to be interpreted as an implausible Theme. In other words, the effect there was caused by the combination of one source of information (from the verb) with another (from the discourse context). In Experiment 1 the bread might well have been anticipated on the basis that it was some butter that was being spread (in other words, given the assignment of one of the verb's available thematic roles to the butter, although an animate Goal given the butter as Theme is by no means ruled out; Bertolucci, 1972), but the observed results do not necessitate that: in the general case, a slice of bread may be a more likely Goal for a verb such as *spread* than a man. Thus, the verb semantics *alone* may enable the anticipation of one object or another as the Goal. It is interesting in this regard to note that in the Altmann and Kamide (1999) studies (and Experiment 2), the post-verbal Theme (the cake) was anticipated during the verb (*eat*). In contrast, we observed in this study no evidence of anticipatory eye movements towards the Goal (the bread or the man) during the verb region (during *spread/slide* or even during *spread/slide the_*). The data suggest, therefore, that when the verb is encountered, the processor does not simultaneously anticipate all possible arguments; perhaps it only anticipates whatever will occur next (in the 3-place case, the next thing after the verb will tend to be the Theme), in which case the second post-verbal argument will only be anticipated once the first post-verbal argument has been encountered. Alternatively, and within a constraint-satisfaction framework, perhaps the processor anticipated arguments (or their referents in the visual world) as a function of the strength of the constraints that had accrued thus far—perhaps the verb alone, unlike the verb plus its first post-verbal argument, did not in fact provide sufficient constraints in Experiment 1 with which to anticipate the identity of the Goal. Further research, beyond the remit of the present series of studies, would be required to establish which of these conclusions (if any) was warranted.

We return to the accrual of constraints in General Discussion. For now, we note that we cannot thus far be sure whether the verb alone, or the combination of that verb with its direct object, led to the anticipatory eye movements we observed in this study. If Experiment 1 had contrasted *The woman will spread the butter on the bread.* with *The woman will spread the plaster on the wall*, the anticipation of the correct Goal *would* have necessitated the combination of the verb with its direct object, so long as one could rule out the possibility that anticipatory looks towards the bread or the wall were not due to some form of direct association between, for example, plaster and walls, or bread and butter (in which case anticipatory looks to the correct Goal might be independent of the verb semantics, and dependent once again on just a single source of information). Such a design would have the additional drawback of entailing a comparison of looks towards different objects during the critical post-verbal region, and we have already indicated some of the problems inherent with such analyses. Nonetheless, Chambers, Eberhard, Filip, and Carlson (2002) reported recently an equivalent study, which avoided some of these problems. In their study, participants heard instructions such as *Pick up the cube. Now put it inside the can.* The visual context contained two cans of different sizes, and depending on the condition, participants either had a (small) cube that fit both cans or a (large) cube that was too big for one of them. In the former case, the eye movements were initially directed towards both cans with roughly equivalent timing, but in the latter case, eye movements were directed sooner to the can that could accommodate the cube. In principle, this study suggests that it *is* the combination of the verb and its direct object (*put* and *it*) that, in these 3-place constructions, gives rise to anticipatory eye movements. However, Chambers et al. (2002) did not report the timing of any saccadic movements

prior to the onset of the target noun (*can*), nor can we be sure that the verb (as opposed to the post-verbal preposition) played any part in the process. Thus, we cannot know from that study whether there *were* anticipatory eye movements as a function of cube size, nor whether they might have resulted from the combination of the verb and the direct object.

It was with these issues in mind that Experiment 2 was devised; to establish whether the effects we have observed in Experiment 1 and previously in Altmann and Kamide (1999) were indeed driven by lexical information associated with the verb alone, or whether such effects could instead (or also) be due to the combination of distinct sources of information.

Experiment 2

Experiment 1 and the Altmann and Kamide (1999) study explored whether one or other of a verb's internal arguments can be anticipated on the basis of information accessed at the verb. Experiment 2 aims to extend these basic findings: In principle, the anticipatory effects reported in Altmann and Kamide (1999) and observed also in Experiment 1 here could have been driven solely by lexical information retrieved at the verb. Alternatively, information from the verb could be combined with information about its grammatical subject to also drive the predictive process. Thus, when the verb in *The boy will eat the cake.* is found to cause anticipatory eye movements towards the cake, is this because cakes are edible things, or because (a) they are edible and (b) they are plausibly eaten by *the boy*? In principle, the predictive process could be driven only by the verb if predictive processes are head-driven and localized to head information only. Experiment 2 attempts to tease apart the two alternatives, by focusing on whether information about the Agent can combine with the selectional restrictions of the verb to predict the Theme: a sequence in which the Agent is a human might predict a different class of food to be the Theme compared to a sequence in which the Agent is a non-human animal (e.g., *The boy ate...* vs. *The cat ate...*). Thus we ask whether the arguments of a verb can be predicted on the basis of *combinatory information* derived from the semantics of the Agent in combination with the verb's selectional restrictions over its post-verbal arguments. The experimental materials of Experiment 2 contain the following auditory sentential conditions (3) and (4) presented in the context of a fairground scene portraying a man, a young girl, a motorbike, a carousel, and various other objects (see Fig. 3, panels A and B indicates the critical regions for data analyses as we shall discuss below).

(3) The man will ride the motorbike.
(4) The girl will ride the carousel.

Both Themes (*motorbike* and *carousel*) satisfy the selectional restrictions of the verb 'ride'. Also, both Agents (*man* and *girl*) are qualified to be an Agent of 'ride': they are both animate. However, knowledge based on real world plausibility suggests that a man would be more likely to ride a motorbike than a carousel while the little girl would be more likely to ride a carousel than a motorbike. Therefore, if combinatory information guides anticipatory eye movements, the motorbike should be looked at more often, before the onset of *motorbike*, in (3) than in (4), whereas the carousel should be looked at before the onset of *carousel* more often in (4) than in (3).

One complicating factor in the design of Experiment 2 is whether anticipatory eye movements to the 'appropriate' Theme objects could be driven by *low-level associations* between the Agent and the Theme object. For example, in the sentence *The cat will eat the mouse.*, anticipatory eye movements to the mouse could simply reflect an association between the lexical item *cat* and mice, independently of the application of the selectional restrictions of the verb *eat*. Such associations need not even be lexically mediated. For example, looks to the cat may cause conceptual information about cats to be accessed that in turn drive looks to the mouse. Or, if the man who will ride the motorbike is portrayed in the scene as wearing a motorcycle helmet (see Fig. 3, panel A), then that information, which is independent of the lexical representation for *man*, might drive eye movements towards the motorbike. To rule out the possibility that anticipatory eye movements in Experiment 2 might be guided by such associations (lexical or conceptual), the experimental material set also includes the following conditions.

(5) The man will taste the beer.
(6) The girl will taste the sweets. [*sweets = candy in British English*].

The verb *taste* in (5) and (6) does not permit either of the Theme objects from (3) and (4); one cannot taste a motorbike or a carousel, and conversely, one cannot ride a beer or ride sweets. If anticipatory eye movements towards the motorbike rather than to the carousel at *ride* in *The man will ride...* are due solely to the man (either the visual entity or its referring expression) and are independent of the verb, we should also observe such movements at the verb in *The man will taste...* (and of course, the converse is true also). As we shall see, this design in fact allows us to factor out the independent contributions of (a) combinatory information, (b) information about the agent, and (c) verb selectional restrictions.

Method

Participants

Sixty-four subjects from the University of York student community took part in this study. They

participated either for course credit or for £2.00. All were native speakers of English and either had uncorrected vision or wore soft contact lenses or spectacles.

Stimuli

Twenty-four scenes (e.g., Fig. 3, panel A) each accompanied by four sentential conditions were constructed (repeated as follows for readers' convenience).

(3) The man will ride the motorbike.
(4) The girl will ride the carousel.
(5) The man will taste the beer.
(6) The girl will taste the sweets.

There were in fact two versions of each visual scene, depending on the Agent of the accompanying sentence. In one version, the man (to use this example) was in one location, and the girl in another, and this version accompanied the sentences in which the man was the Agent. For the sentences where the girl was the Agent, the original version of the scene was changed by swapping the positions of the man and the girl (and where necessary, adjusting their sizes to take perspective into account). This ensured that the Agent of the event described in the sentence would always occupy the same spatial position relative to the other objects in the scene (thus ruling out biases in the subsequent eye-tracking experiment that could arise if the relative positions changed across the sentence types). Occasionally, this also meant that we had to left–right reverse one of the Agents to ensure that in both sentences the Agent faced in approximately the same direction.

The auditory stimuli were recorded and edited as in the previous experiment. In cases of particle verbs such as 'climb up,' verb offset was marked as the offset of the particle.[2] Twenty-four unrelated filler-items were included, consisting of both 2-place (monotransitive) and 3-place constructions. The stimuli were presented in a fixed-random order such that no adjacent items belonged to the same experimental condition. The experimental sentences for Experiment 2 are given in Appendix B.

Procedure

The same procedure was used as in Experiment 1, except that the onset of each visual scene preceded sentence onset by 1000 ms (the duration of this preview was motivated by the complexity of the visual scenes relative to Experiment 1, and was determined by the subjective judgment of the authors).[3]

Analyses

Our primary prediction is that looks towards the motorbike will be facilitated, at the verb, by the *combination* of lexical information associated with both the verb and its preceding subject, and that in line with the hypothesis that constraints are applied *as they accrue* (cf. Altmann & Steedman, 1988; MacDonald et al., 1994), we should observe, at the verb, no independent influence on looks towards the motorbike due just to the sentential subject or just to the verb's selectional restrictions. Consequently, our hypothesis necessitates planned comparisons to establish that (a) there will be more looks to the motorbike after *The man will ride* than after either *The girl will ride*, *The man will taste*, or *The girl will taste*, and (b) these last three cases do not differ (if they did, that would suggest independent influences of the subject and/or the verb's selectional restrictions). The comparisons are described in further detail below.

Combinatory effects. To explore the combinatory effect, one might assume at first glance that one should compare looks towards the motorbike after *The man will ride* with looks towards the carousel. However, and as stated earlier, this study was designed in such a way as to permit comparisons across conditions for looks to the same object. Thus, the corresponding comparison would be between looks towards the motorbike after *The man will ride* and looks towards the motorbike after *The girl will ride* (if only verb-based information drove the anticipatory process, there would be no difference between these two conditions). However, given the potential for an 'agent effect' (see above), one should compare also looks towards the motorbike after *The man will ride* with looks towards the motorbike after *The man will taste* (an agent-only effect in the absence of a verb effect would predict no difference between these two conditions). Consequently, evidence of a combinatory effect in the absence of any other effect requires both these comparisons. The first comparison would in fact comprise a comparison of the average of looks to the motorbike after

[2] There were four experimental sentences (out of 48) that contained a pre-position before the onset of the second NP (*climb up*, *jump through*, *sleep in*, *swim in*). In those cases, it would be more accurate to refer to the NP2 as the Location/Goal than the Theme. However, for simplicity, we overgeneralize the thematic role of NP2 for the rest of our discussion.

[3] We note that whereas in Experiment 1 (*The woman will spread/slide the butter on/to…*), participants launched saccadic eye movements towards the Agent of the sentence during *The woman will* on around 85% of trials, they did so in Experiment 2 (during *The man will* or *The girl will*) on only around 56% of trials. We attribute this difference to the 1000 ms preview that participants received in Experiment 2. Interestingly participants launched a look towards the butter during the post-verbal *butter* on 32% of trials in Experiment 1, and launched a look towards the motorbike during the post-verbal *motorbike* on 32% of trials in Experiment 2. Evidently, further research is required to establish the influence of preview on data such as these.

The man will ride, looks to the beer after *The man will taste*, looks to the carousel after *The girl will ride*, and looks to the sweets after *The girl will taste* with the average of looks to the motorbike after *The girl will ride*, looks to the beer after *The girl will taste*, looks to the carousel after *The man will ride*, and looks to the sweets after *The man will taste*.

Agent effects. To rule out (or rule in) an effect due solely to the identity of the agent, we compare looks to the motorbike after *The man will taste* with looks to the motorbike after *The girl will taste*—an agent-only effect would predict more looks to the motorbike in response to the former than in response to the latter (or it would predict more looks to the beer after *the man will ride* than after *the girl will ride*; to the carousel after *The girl will taste* than after *The man will taste*, and so on).

Verb effects. To rule out (or rule in) an effect due solely to the selection restrictions associated with the verb, we compare looks to the motorbike after *The girl will ride* with looks to the motorbike after *The girl will taste*—selectional restrictions would favor looks to the motorbike in the former condition (and looks to the carousel after *The man will ride* than after *The man will taste*, and so on).

Each of these comparisons entails averaging looks to the four different objects in the four appropriate sentential conditions. These comparisons are summarized in Table 1. In our discussion of the data, we shall refer to, for example, the motorbike being looked at more after *The man will ride* than after *The girl will ride*, but this is just shorthand for saying that, in addition, the sweets were looked at more after *The girl will eat* than after *The man will eat*, the carousel more after *The girl will ride* than after *The man will ride*, and the beer more after *The man will taste* than after *The girl will taste*—each data point to be reported will be the average of all these cases.

Results

The eye movement data were analysed using the same procedures adopted for Experiment 1. We sought to determine when, relative to the onset of the Theme noun (*motorbike*, *carousel*, etc.), the participants' eyes first moved to the target region in the accompanying visual scene. For the data analyses, we divided the relevant data points into two regions: *Region 1*, verb onset to verb offset (*ride/taste*); *Region 2*, verb offset to Theme onset (*_the_*). Fig. 3, panel B illustrates this division, and the full data across all regions are graphed in Fig. 4.

The overall mean duration for each region was as follows: Region 1—408 ms; Region 2—229 ms. Given the 200 ms or so to plan and initiate an eye movement, the second region is essentially arbitrary, insofar as saccadic movements during that region most likely do not reflect processes initiated within that region. As we shall shortly see, the data from this region are informative nonetheless. For each region, we calculated the percentage of trials in which saccades were launched towards each object. Saccadic movements whose onset was prior to the region onset were eliminated from the analyses for that region. For example, any saccadic movements starting before the verb were not included in the data for the verb region.

Region 1: *Ride/taste*

Overall, looks were initiated during the verb towards the target objects (the motorbike, the carousel, the beer, and the sweets) on 25% of trials. Looks were initiated towards the Agent on 19% of trials, and to the other person on 7% of trials. Given the duration of this region (408 ms) there were virtually no trials in which there was more than one fixation. Thus, on approximately 50% of trials a look was initiated during the verb to one of the critical objects or to one of the two animate entities.

In respect of the crucial comparisons across conditions, all statistical analyses were performed on the arcsine-transformed proportions of trials with at least one look directed at the target object. All comparisons were based on the error term from a one-way analysis of variance with four levels corresponding to the four conditions outlined above (the effect of condition was significant on both the subject and item analysis: $F1(3,189) = 3.3,\ p < .03;\ F2(3,69) = 3.6,\ p < .02$). For all analyses (in this region and the next), exactly the same statistical patterns were observed using paired *t*-tests on each pairwise comparison.

Combinatory effects. There were more looks to the motorbike in the *The man will ride* condition than in the *The girl will ride* condition (9 vs. 6%: $F1(1,189) = 7.3$,

Table 1
Summary of comparisons used to explore the Combinatory, Agent, and Verb effects in Experiment 2

	Motorbike	Carousel	Beer	Sweets
Combinatory effect (I)	*man ride* vs. *girl ride*	*girl ride* vs. *man ride*	*man taste* vs. *girl taste*	*girl taste* vs. *man taste*
Combinatory effect (II)	*man ride* vs. *man taste*	*girl ride* vs. *girl taste*	*man taste* vs. *man ride*	*girl taste* vs. *girl ride*
Agent effect	*man taste* vs. *girl taste*	*girl taste* vs. *man taste*	*man ride* vs. *girl ride*	*girl ride* vs. *man ride*
Verb effect	*girl ride* vs. *girl taste*	*man ride* vs. *man taste*	*girl taste* vs. *girl ride*	*man taste* vs. *man ride*

The table should be read as follows: to explore the Agent effect, an average was calculated of looks to the motorbike in the *man taste* condition, looks to the carousel in the *girl taste* condition, looks to the beer in the *man ride* condition, and looks to the sweets in the *girl ride* condition. This average was compared against the average of looks to the motorbike in the *girl taste* condition, looks to the carousel in the *man taste* condition, looks to the beer in the *girl ride* condition, and looks to the sweets in the *man ride* condition.

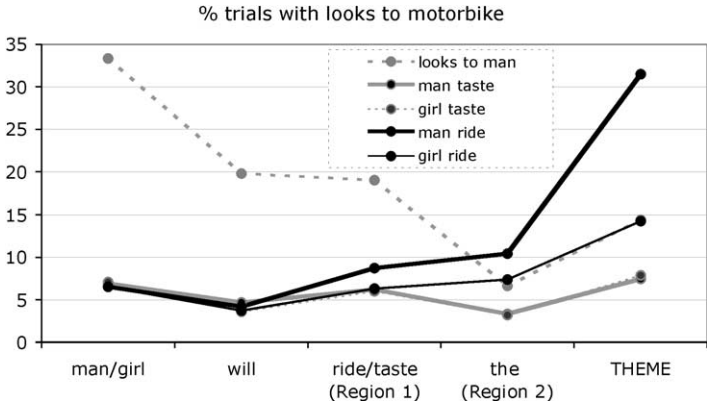

Fig. 4. Percentage of trials for each region of interest with looks to the motorbike in each of the four different sentential conditions (each datapoint include looks to the other objects, averaged across the appropriate conditions—see the main text). Also included are looks to the Agent.

$p < .008; F2(1,69) = 6.6,\ p < .02$), and in the *The man will ride* condition than in the *The man will taste* condition, although the effect does just miss significance in the by-subjects analysis (9 vs. 6%: $F1(1,189) = 3.8,\ p < .052;\ F2(1,69) = 6.4,\ p < .01$).

Agent effects. Looks to the motorbike did not differ between the *The man will taste* condition and the *The girl will taste* condition (both 6%: both $F < 1$).

Verb effects. Looks to the motorbike did not differ across the *The girl will ride* condition and the *The girl will taste* condition (both 6%: both $F < 1$).

Region 2: '*_the_*'

Overall, looks were initiated during '*_the_*' towards the target objects (the motorbike, the carousel, the beer, and the sweets) on 24% of trials. Looks were initiated towards the Agent on 7% of trials, and to the other person on 3% of trials. Thus, on approximately 35% of trials a look was initiated during this short region (229 ms) to one of the critical objects or to one of the two animate entities. The effect of condition was significant on both the subject and item analysis: $F1(3,189) = 27.2,\ p < .0001;\ F2(3,69) = 18.8,\ p < .0001$.

Combinatory effects. There were more looks to the motorbike in the *The man will ride* condition than in the *The girl will ride* condition (10 vs. 7%: $F1(1,189) = 5.6,\ p < .02;\ F2(1,69) = 4.4,\ p = .04$), and in the *The man will ride* condition than in the *The man will taste* condition (10 vs. 3%: $F1(1,189) = 52.0,\ p < .0001;\ F2(1,69) = 35.7,\ p < .0001$). We note that the latter comparison is more robust, statistically, than in the previous region where it just missed significance ($p = .052$) in the by-subjects analysis.

Agent effects. Looks to the motorbike did not differ between the *The man will taste* condition and the *The girl will taste* condition (both 3%: both $F < 1$).

Verb effects. There were more looks to the target object in the *The girl will ride* condition than in the *The girl will taste* condition (7 and 3%. respectively: $F1(1,189) = 26.3,\ p < .0001;\ F2(1,69) = 17.6,\ p < .000$).

Discussion

To summarise the results, combinatory effects were manifest at the verb: increased looks to the motorbike after *The man will ride* were not due simply to the selectional restrictions conveyed by the verb independently of the subject (otherwise there would have been more looks to the motorbike after *The girl will ride* than after *The girl will taste*). Nor were they due to some direct association between the man and the motorbike independently of the verb (otherwise there would have been more looks to the motorbike after *The man will taste* than after *The girl will taste*). Interestingly, in the immediately post-verbal region ('*_the_*'), there *was* an apparent effect of the verb's selectional restrictions independently of the combinatory effect. Most likely, this does not reflect the later application of selectional restrictions (what would be the point, when they are subsumed by the combinatory information?), but reflects instead a 'contrast effect' (cf. Sedivy, Tanenhaus, Chambers, & Carlson, 1999). Sedivy et al. (1999) presented participants with a scene containing a tall glass and a short glass (and various distractors), and participants either had to *Pick up the tall glass* (Experiment 2) or to verify the question *Is there a tall glass?* (Experiment 3). Although looks to the target glass (the tall glass) increased rapidly even before noun onset, looks to the *short* glass (the contrasting object) relative to the unrelated distractors also increased, but a short time later. Sedivy et al. (1999) suggested that this later rise in looks might reflect a process of confirming that the appropriate object, given the contrast, had been selected (so participants would look to the tall glass, but then 'double-check' the short glass). Within our own data, we interpret the late 'verb' effect in a similar manner: the motorbike is

the most plausibly ridden object given the man as agent, but only in the context of the concurrent visual scene, where the contrast is with the carousel. If the Sedivy et al. (1999) contrast effect generalises, we would expect, then, an early rise in looks to the motorbike, but a later rise in looks to the carousel. In terms of looks to the motorbike in the different sentential conditions, this translates into a (late) rise in looks to the motorbike after *The girl will ride* (because of the contrast effect) relative to looks to the motorbike after *The girl will taste* (where the motorbike does not contrast with any other object in respect of being a plausible object for the verb). This is precisely what was found. And as a corollary to this effect, one would also expect more looks to the motorbike in the later region after *The girl will ride* than after *The man will taste* (where again, it does not enter into any contrast set with respect to being a plausible object for the verb), and this was indeed the case (7 vs. 3%, respectively: $F1(1,189) = 23.6$, $p < .0001$; $F2(1,69) = 15.1$, $p = .0002$). Both these differences are apparent in Fig. 4.

One further aspect of the data that warrants discussion is the relatively small number of trials in which there were looks to the motorbike (or equivalent). In fact, in the region '*ride the_*', looks were initiated to one or other of the potential objects on 44% of trials. If these were randomly 'targeted' at each of the objects, we could expect 11% trials with looks to any one object. Again, across the region '*ride the_*' as a whole, looks to the plausible object (the motorbike in the *The man will ride the* condition, for example) occurred on 19% of trials, compared to 13% of trials towards the implausible object ($F1(1,189) = 12.9$, $p = .0004$; $F2(1,69) = 9.8$, $p < .003$). On approximately 35% of trials during this period, looks were launched towards either the Agent (26%) or the other person in the scene (9%). Thus, only around 20% of trials are 'unaccounted for', meaning that on these trials there were no looks launched towards any of the objects in the scene in the region *ride the_*' (in fact, on 23% of trials there were no saccadic movements away from whatever was being fixated at the onset of *ride the_*'—the numbers do not add up to 100% because of the small number of occasions when multiple fixations occurred within this 638 ms. interval). There are two reasons why such trials might occur. First, participants might not launch a movement towards any object in any given time window because they are already looking at an object that, for whatever reason, holds their attention (although in only 7% of trials in which there was no saccadic movement during '*ride the_*' away from whatever was being fixated did that fixated object correspond to the plausible object). Second, eye movements reflect overt shifts in attention, but they do not necessarily reflect shifts in attention within the mental representation of the visual scene. While it is necessarily the case that a saccadic eye movement must be preceded by a shift in attention, it is not necessarily the case that a shift in attention (to some other component of the mental representation) is followed by a saccadic eye movement. In this respect, overt shifts in attention as evidenced by saccadic eye movements underestimate the true extent of attentional shifts (given the possibility of attentional shifts that do not manifest themselves as shifts in overt visual attention).

There is little doubt that the effects we have observed in this experiment, although temporally robust and statistically reliable, are small effects. We cannot claim, therefore, that on 100% of occasions in which a participant hears *The man will ride* they anticipate the motorbike as the object of the riding, and that they do so on the basis of the combination of information from the verb and from its grammatical subject. What we do claim, however, is that the processing system is able, on occasion at least, to form anticipations of this kind. and that it most likely does so on a probabilistic basis (see General Discussion).

The main goal of Experiment 2 was to investigate whether the verb is the only information source able to support the prediction of forthcoming arguments. The potential existence of those arguments is necessarily determined by the verb, but the issue concerns not so much whether those arguments are predicted, but rather what information about them is predicted, and what sources of information underlie the prediction. Experiment 2 demonstrated that the semantic properties of a forthcoming Theme are predicted, and on the basis of the combination of information about the Agent and about the verb. We return in our General Discussion to how such effects might be manifest within a constraint-satisfaction model.

The timing of the effects we observed—localized around verb offset and immediately beyond, suggests rapid integration of verb information with information about the verb's subject, and rapid integration of this combination with real world knowledge regarding the plausibility of the objects in the visual scene as potential Themes given that combination. This begs the question of the extent to which the verb 'drives' the predictive process. Given that the verb encodes information about its argument structure, it could be supposed that the predictive processes identified thus far are verb-specific—that is, a verb's arguments can only be predicted at or after the verb, as it is the verb that encodes the potential for those arguments to be realized in the input. However, in encoding those arguments, the verb provides particularly strong constraints on what may come next; thus, it may not be the category 'verb' per se that is important here, rather the strength of the constraints that accompany lexical items of that type. We explore this directly in Experiment 3.

A second question that is hitherto unresolved is whether the effects we have observed in Experiment 2 are due to the *syntactic combination* of the verb with its subject, and subsequent interpretation of that combination, or whether the effects are due to the combination

of semantic information independently of any syntactic processing. Did that combination entail a structural commitment by the processor in respect of the pre-verbal noun phrase functioning as the subject of the verb? This questions is not unrelated to the question of whether the processor might have projected syntactic structure at the verb that corresponded to the hitherto unencountered post-verbal direct object. Did the processor anticipate the Theme in Experiment 2 on the basis of structural sensitivities, or on the basis of semantic sensitivities? Of course, there may be little difference between the two if semantics is a reflection of the co-occurrence of words (or objects) and their contexts (cf. experiential approaches to semantics; e.g., McRae et al., 1997; and see Altmann, 1997). Experiment 2 does not address whether the basis of the anticipatory behavior observed there was structural as well as semantic.

In Experiment 3, we address directly the issues of structural sensitivity and verb-specificity. We turn to the processing of the verb-final language Japanese and explore whether, in 3-NP sentences, the syntactic case-marking on the first two noun phrases (NP1: subject; NP2: indirect object) can be used to anticipate the semantic properties of whatever might subsequently occur as a third noun phrase (NP3: direct object) in the sequence NP1–NP2–NP3–verb. If NP1–NP2 combinations can predict NP3 in Japanese, we would have direct evidence that the prediction of a verb's arguments do not necessitate the verb itself. Moreover, if the case-marking on the NP1–NP2 combination determines the prediction, we would have direct evidence that syntactic information interacts with real-world knowledge during the formulation of the prediction.

Experiment 3

In Japanese, all arguments of the verb appear prior to the verb. For example, consider the following example:

(7) 少女が猫に魚をやった。
syoojo-ga neko-ni sakana-o yatta.
girl-nominative cat-dative fish-accusative gave.
The girl gave the fish to the cat.

In (7), all three of the verb's arguments come before the verb, and as is typical with head-final languages, each argument is case-marked (in Japanese, the case-marking is not a nominal inflection but rather a post-nominal particle). If verb's are the driving force behind predictive processing, there should not be any opportunity in Japanese constructions such as this for predictive processing of forthcoming arguments. In contrast, it is nonetheless possible that the processor can use information appearing prior to the verb in order to predict properties of forthcoming items: For example, the information extracted from '*girl*-nom' and '*cat*-dat' (henceforth, we abbreviate the case-markers) suggests that it is likely that the girl serves as the Agent and the cat serves as the Goal in the event being described by the sentence. Two sources of information can then be used to predict properties of an upcoming Theme: first, the vast majority of sentences in Japanese beginning 'NP1-nominative NP2-dative' signal a 3-NP construction in which an accusative NP3 will follow prior to the sentence-final verb. Consequently, the processor can anticipate a verb that denotes an act of transference from the Agent (manifested linguistically as NP1-nominative) to the Goal (NP2-dative), with the subsequent Theme (anticipated as NP3-dative) being whatever was transferred. Knowledge of the real world contingencies that involve cats might enable the processor to predict further that whatever will appear as Theme will be something plausibly transferred to cats by girls.

Experiment 3 uses the visual-world paradigm to explore pre-head (i.e., pre-verb) prediction in Japanese. An example material set to accompany the visual scene shown in Fig. 5, panel A is as follows (Fig. 5, panel B indicates the critical regions for data analyses as we shall discuss below).

(8) *Dative condition.*
ウェイトレスが客に楽し気にハンバーガーを運ぶ。
weitoresu-ga kyaku-ni tanosigeni hanbaagaa-o hakobu.
waitress-nom customer-dat merrily hamburger-acc bring.
The waitress will merrily bring the hamburger to the customer.

(9) *Accusative condition.*
ウェイトレスが客を楽し気にからかう。
weitoresu-ga kyaku-o tanosigeni karakau.
waitress-nom customer-acc merrily tease.
The waitress will merrily tease the customer.

At issue is whether, after hearing '*waitress*-nom *customer*-dat' in the Dative condition (8) above, the processor will anticipate that the third argument will refer to the one object in the scene that could plausibly be transferred by the waitress to the customer—namely, the hamburger. The comparison with the Accusative condition (9) above is crucial: In the Dative condition, NP-acc is obligatory after the sequence 'NP1-nom NP2-dat' unless the verb turns out to be one of the very few monotransitive verbs that take an NP-dat as its Theme (e.g., *aisatu-suru*—'greet'), or the NP-dat is interpreted as a Goal after an intransitive verb (e.g., 'iku'—go). However, in the Accusative condition, there are two alternative ways in which the fragment 'NP1-nom NP2-acc' can continue—either as part of a monotransitive construction, in which case there will be no further arguments, or as part of a 3-NP construction. In the latter case, there scene contains no plausible object that could be referred to as Goal (neither the hamburger nor the dustbin are plausible Goals). Thus, if the first two noun phrases provide enough information to enable the pre-

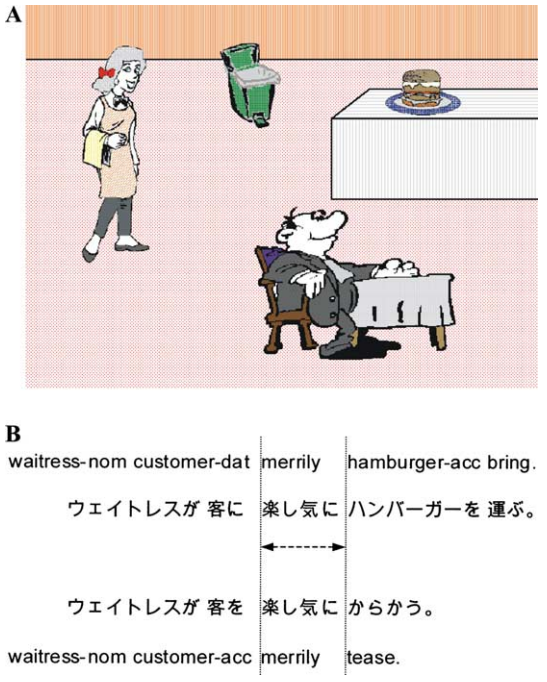

Fig. 5. An example visual (panel A) and auditory (panel B) stimuli from Experiment 1. Both the Japanese scripts and glosses are provided for the auditory sentences. The region relevant to the analyses is indicated with an arrow. The relative sizes of the space between the words do not correspond to the actual relative lengths of the pauses in the sentences.

diction of forthcoming material, there should be more anticipatory looks towards the hamburger in the Dative condition, when it can plausibly fulfill the Theme function, than in the Accusative condition, when it cannot plausibly fulfill any function that might be anticipated on the basis of the sequence 'NP1-nom NP2-acc'. Importantly, any difference in anticipatory looks could only be attributed to the difference in the case-marking of NP2—suggesting that syntactic information can itself drive the predictive process, and that the predictive operations we have observed in our previous studies need not be attributed solely to semantic processes operating independently of the establishment of syntactic dependencies during parsing.

Method

Participants

Twenty-two participants from the University of York student community took part in this study. They were paid £2.00 for their participation. All were native speakers of Japanese and the average duration of having lived in an English-speaking country was 20.45 months. They all either had uncorrected vision or wore contact lenses or spectacles.

Stimuli

Sixteen sets of experimental items were created (see examples 8 and 9 above, and Fig. 5). Out of the 16, 8 sets had an animate noun in the second argument position as shown in the previous examples, and the other 8 had an inanimate noun in NP2 position (e.g., *man*-nom *wall*-dat *carefully poster*-acc *hung*/*man*-nom *wall*-acc *carefully wipe*—where the dative indicates a location role). The visual scenes were created as in the previous experiments. The auditory stimuli were recorded by a female native speaker of Japanese in the standard accent (a Tokyo accent). Additionally, a further 16 foil sentences like the following were included:

(18) 医者がぬいぐるみをやさしく子供に与える。
isya-ga nuigurumi-o yasasiku kodomo-ni ataeru.
doctor-nom soft toy-acc gently child-dat give.
The doctor will gently give the soft toy to the child.

The foil sentences had an NP-acc in the second argument position, followed by an adverb, an NP-dat, and the verb. This scrambled structure was included to prevent participants from learning that the NP-acc would always end in a monotransitive structure. Sixteen further items were included as fillers. Half of these fillers had a structure 'NP-nom NP-with adverb NP-acc verb,' and the other half 'NP-nom NP-post-position (next to/above/beside) adverb NP-acc verb.' Altogether, there were two subject groups, and each subject heard 8 Dative condition sentences, 8 Accusative condition sentences, 16 Foils, and 16 Fillers (48 trials overall). The materials were arranged in a fixed-random order such that no adjacent items belonged to the same experimental condition. The experimental sentences for Experiment 3 are given in Appendix C. There were two experimental stimuli lists, such that each material set appeared in both conditions, but no material set appeared more than once in a given list.

Procedure

The same procedure was used as in Experiment 1. The onset of the visual scene coincided with the onset of the spoken sentence, and the trial was terminated after 8 s. All written and oral instructions were given in Japanese.

Results

The eye movement data were treated in the same way as in Experiments 1 and 2 for the analyses. The arcsine-transformed data for looks to the target object between the onset of the adverb (*merrily*) and the onset of the fourth constituent (*hamburger*-nom' or *tease*) were entered into the analyses (see Fig. 5, panel B). As before, the data consisted of the percentage of trials in which at least one saccadic eye movement was initiated to the target object within the critical period (mean duration of

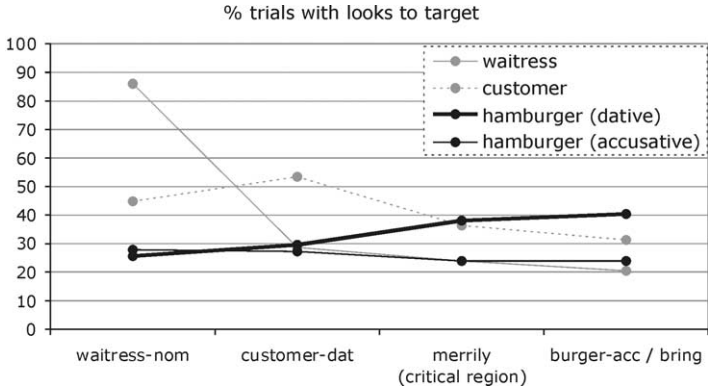

Fig. 6. Percentage of trials for each region of interest with looks to the hamburger in each of the two different sentential conditions.

the region was 872 ms). The data for the different word positions are graphed in Fig. 6.

There were significantly more looks to the Target objects in the Dative condition (38.1%) than in the Accusative condition (23.9%) during the critical period ($F1(21) = 7.0,\ p < .02;\ F2(15) = 7.5,\ p < .02$). The pattern of fixations, therefore, suggests that people anticipated which object in the picture would be most likely to follow as the most plausible Theme before hearing the onset of the corresponding referring expression (NP3).

Discussion

Experiment 3 aimed to investigate the role of the grammatical head in the prediction of its arguments. Experiments 1 and 2 explored the extent to which verb-based information can be used to predict a forthcoming Theme in English active constructions. Experiment 3, using Japanese verb-final constructions clearly showed anticipatory eye movements towards the potential referent of NP3 when only NP1 and NP2 had been encountered: after encountering the sequence 'NP1 NP2', participants looked at the most plausible Theme object in the picture, during the subsequent adverb (but before the onset of NP3) more often when NP2 was case-marked as dative (requiring a subsequent Theme) than when NP2 was case-marked as accusative. Hence, the experiment demonstrated that prediction of forthcoming arguments is possible even in the absence of the grammatical head. Moreover, given that the only difference between the Dative and Accusative conditions was in the case-marking ('NP1-nom NP2-dat adv' vs. 'NP1-nom NP2-acc adv'), we conclude that this prediction was in part based on syntactic information regarding case-structure in Japanese, and not solely on interpretive processes operating independently of syntax (the use of case-marking information in prediction has been also found in a visual-world experiment using German in Kamide, Scheepers, & Altmann, 2003).[4]

Our evidence for pre-head prediction is relevant for the debate as to whether or not the parser is head-driven (see Kamide & Mitchell, 1999, for a review); prediction of the kind we have observed, prior to the grammatical head, would be impossible without some structural analysis of the existing pre-head arguments—we take it as given that if the processor's operations are contingent on particular case-markers, those operations can be deemed to be sensitive to structural analysis of the input. In principle, however, the data we have reported here need not have arisen through sensitivity to *sequences* of case-marked noun phrases; perhaps the processor can determine on the basis of *customer*-dative', but independently of any knowledge about the Agent (the waitress), that something is likely to be referred to that is a plausible thing for the customer to receive. To demonstrate the use of combinatory information across different noun phrases would require a design similar to that of Experiment 2: contrasting, for example, *waitress*-nominative *customer*-dative *hamburger*-accusative *bring* with *pharmacist*-nominative *customer*-dative *medicine*-accusative *bring*. In the absence, currently, of such a design, our conclusions must be limited to the observation that the processor is able to anticipate subsequent reference (or the inclusion of a visual object into the event being described) on the

[4] It is unlikely that these data are due to the anticipatory coarticulation of the post-adverbial segment on the adverb itself. First, there was on average a 145 ms. delay between the offset of the adverb and the onset of the next segment. Second, the adverbs were relatively long, consisting on average of 5.1 morae. If anticipatory coarticulatory information from the following word did appear on the adverb, it most likely would affect only the last mora, in which case it occurred too late to influence the pattern of eye movements measured during the lifetime of the adverb (given the time it takes to launch a saccadic eye movement).

basis of the combination of case-marking and lexico-semantic information. We do believe, nonetheless, that the data from Experiment 2 converge on the hypothesis that in the Japanese case also, we would most likely find that the processor *is* sensitive to sequences of case-marked nouns, and is able to use combinatory information to drive the anticipatory process. We return in General Discussion to whether the processor in fact anticipates *reference*, or simply assumes that a particular object in the scene most likely takes part in the event described so far.

Arguably, the visual-world paradigm as used here offers a relatively more direct method for investigating the processing of pre-head arguments than traditionally used methods. For example, some previous studies have investigated the reading of complex sentences and have taken prolonged reading times in particular regions of the sentences as evidence for (incorrect) structural attachments made prior to the head (e.g., Bader & Lasser, 1994; Koh, 1997; Kamide & Mitchell, 1999). In other studies, the additional cost of processing non-canonically-ordered (or *scrambled*) pre-head arguments (relative to canonical orderings) has also been taken as demonstration of pre-head attachment of the arguments (e.g., Hemforth, 1993; Mazuka, Itoh, & Kondo, 2000; Miyamoto & Takahashi, 2002). In short, these other demonstrations of pre-head attachment have all relied on indirect indicators of pre-head structural interpretation. In contrast, the experimental technique and manipulation we adopted in the present experiment enabled us to examine directly the consequences of pre-head processing on the guidance of overt visual attention, and to establish a direct link between syntactic structure and semantic interpretation during pre-head processing.

Experiment 3 also offers some implications (albeit rather speculative) on various other important issues in Japanese sentence processing. For example, it indicates that, given the particular visual contexts, analyzing 'NP-dat' (e.g., *customer*-dat) as the Goal of a 3-place verb is preferred over analyzing it as the Theme or Goal of a monotransitive verb (these other analyses would not have engendered more looks towards the target object relative to the monotransitive accusative condition). As mentioned above, such constructions are relatively rare in Japanese, and one interpretation of our data is that this preference is constrained by information regarding the frequency of such structures. However, an alternative account of our data would hold that the preference for analyzing 'NP-dat' as the Goal of a 3-place verb arose because of the availability in the concurrent visual scene of an object that could plausibly take on the Theme role. We cannot distinguish on the basis of the present data whether the behavior we observed was based on such pragmatic affordances (that is, that the context itself provided a cue as to which structure was likely) or whether it was based on knowledge of the statistical distributions of particular verb types in Japanese. However, this issue is incidental to the main aims of the study, which were to demonstrate pre-head prediction of forthcoming arguments in a head-final language.[5] We return below to consideration of the extent to which the visual context itself may influence the interpretation of the concurrent linguistic structure, and the extent to which objects within that context may be interpreted as playing a role in the event described by that structure, but in the absence of any linguistic projection, or prediction.

General discussion

The primary issue in the present paper has been whether preceding linguistic input can provide the basis for the processor to anticipate upcoming input, and if so, what kinds of information provided by that preceding input provide that basis. Previously, we have shown that selectional restrictions can be used to predict the semantic properties of whatever will be referred to post-verbally (with the assumption that the processor predicts that something *will* be referred to post-verbally on the basis of the verb's argument structure—however, even this is a moot point, to which we return below). In this paper, we have explored further the information on which such predictions are based.

Experiment 1 investigated three questions: (a) whether post-verbal arguments in English other than Themes—Goals—can be predicted; (b) whether anticipatory eye movements are restricted to portions of the linguistic input that are not overtly referential, or whe-

[5] A related and equally incidental issue concerns the processing of *NP1-nom NP2-acc* sequences and whether a subsequent optional and non-canonically ordered Goal argument is projected. The issues are related insofar as the non-canonical ordering is the less frequent, and again, the scenes deployed in Experiment 3 did not provide any plausible Goal objects. To explore this issue further one could include a potential Goal for the Accusative condition in the picture, providing the degree of the plausibility of the events in both conditions is equal. For example, one material set would include *woman*-nom *bread*-dat *sluggishly butter*-acc *spread*.' (Dative condition) and *woman*-nom *bread*-acc sluggishly *plate*-dat *put*.' (Accusative condition), in the context of a scene with a woman, some bread, some butter, a plate, and a coffee cup. Then, supposing that the woman spreading the butter on the bread is as plausible an action as is the woman putting the bread on the plate, we might expect to see more anticipatory eye movements to referents for the most plausible third argument in each condition (the butter in the Dative condition, and the plate in the Accusative condition). This would demonstrate projection of both an obligatory Theme after *Agent + Goal* and an optional Goal after *Agent + Theme*. Such further studies have been planned and started in our own lab in Japanese (partial results have been reported in Kamide & Altmann, 2000), and also in collaboration with Christoph Scheepers and Matt Crocker in German verb-final structures.

ther such movements towards one object in the scene can be found during an expression that refers to a different object in the scene; and (c) whether the processes that underpin this anticipatory behavior are restricted to tasks in which participants have to, or are encouraged to 'verify' a sentence against a matching scene, or whether they can occur as part of a 'look and listen' task in which all sentences match their corresponding scenes. In the event, anticipatory eye movements to the most plausible Goal object in the scene were observed during the lifetime of the preceding referring expression (that is, the expression preceding the one that would be assigned the Goal role). The remaining two studies focused on more theoretical aspects of prediction. Experiment 2 found that semantic (real world) properties of a verb's Theme can be anticipated on the basis of the combination of that verb's semantic information with properties of the Agent, while Experiment 3 found that properties of the Theme can be anticipated on the basis of the combination of lexico-semantic and case-marking information in a head-final language (in which the involvement of the verb is ruled out). Whereas Experiment 3 did not unambiguously implicate combinatory information derived across distinct lexical items as the basis for such anticipation, Experiment 2 did precisely that.

In order to consider the nature of the processes that underpinned these effects, we need to distinguish between two kinds of anticipation, only one of which entails prediction. Thus far, we have used the terms interchangeably, but there is a distinction that can be drawn between these terms in respect of the content of the representations that drive the anticipatory eye movements we have observed in these and other studies. 'Prediction' entails the explicit representation of something that will be true in the future, and thus constitutes a projection 'forward in time' of something that may not be true now but that is assumed to become true (possibly at some specified time) with some probability that can be derived from the current evidence. But there is a form of anticipation in which something can be assumed to be true, again with some probability on the basis of the current evidence, but without entailing any representation 'forward in time'. The distinction can be made most easily in the context of Experiment 2 (*The man/girl will ride...*). Anticipatory eye movements to the appropriate object in the accompanying scene (the motorbike or the carousel) could indicate that the processor predicted—that is, projected forward from the verb—a post-verbal argument corresponding to that verb's grammatical object. This projected structure would include grammatical information (it must be a noun phrase and possibly, in the Japanese case, have a particular case-marking), and perhaps semantic information also (the referent must satisfy the selectional constraints associated with the verb). This projection would constitute a prediction that a certain kind of linguistic input (with, perhaps, a certain form) would be forthcoming. The projected structure could then be evaluated against the visual context or its mental representation (that is, a search made for a suitable referent for the postulated noun phrase), resulting in the observed shift of attention onto the appropriate referent. Alternatively, the anticipatory eye movements could be driven instead (or more probably also) by processes that are not language-specific but which attempt to organize thematically the objects in the visual context, both with respect to each other and with respect to whatever representations are available (e.g., from the language) that may constrain their thematic interpretation. In effect, then, the processing system (linguistic and/or non-linguistic) assumes that some subset of the objects in the visual context will take part in the event being described linguistically, and attempts to find roles for those objects as that description unfolds and as those roles become available. The verb *ride* makes available a Theme role, and the processor evaluates each object in the visual context (or its mental representation) against the constraints on whatever should fulfill this role. In this case, such constraints include that the object should be rideable and plausibly ridden by whoever the subject of the sentence refers to. Hence the shift of attention towards the appropriate object. In this version of anticipation, there is no projection of structure forward in time; something is caused to happen before it otherwise might, but without any explicit prediction that it might happen or of *when* it might happen—assumptions are made regarding thematic role assignments (or these thematic role assignments are actually made) without making any assumptions at all regarding subsequent input.

What we do not know, then, is whether for a fragment such as *The man will ride...*, the processor activates representations at *ride* that correspond to the range of rideable objects that a man might ride *even in the absence* of a concurrent visual context portraying one or more rideable things. Thus, the paradigm does not allow us to determine whether it is the linguistic structure that triggers a predictive process, because the syntax determines that something is about to be referred to, or whether it is the visual context that suggests to the processor that something might plausibly enter into a thematic relationship with the man, mediated by the act of riding. We do not, needless to say, have data that empirically address this distinction one way or the other. We do note, however, that there is something of an equivalence between the two alternatives presented above: the verb *ride* makes available a Theme role only insofar as this is a part of its lexical specification. But that specification must in part be experiential; if *ride* did not occur in the context of a subsequent grammatical object that, across different experiences, shared certain semantic characteristics, its lexical specification would not reflect the particular argument structure that it does, and nor would it reflect

the semantic constraints on the arguments that can fill that structure. Thus, the lexical specification implicitly assumes a predictive component (a verb's argument structure is, after all, a reflection of what may have come before and what may come after). In some grammatical theories such a component is made quite explicit (cf. Combinatory Categorical Grammar; e.g., Steedman, 2000). Whether the anticipatory eye movements we have observed reflect the projection of structure (syntactic or otherwise) that is explicitly tagged as representing future input, or whether it does not, is less relevant to the present discussion than the fact that thematic dependencies can be computed in advance of their corresponding grammatical dependencies. We are more concerned with the information, or evidence, that enables this computation. This said, we are necessarily concerned with the content of the representations constructed by such processes. In the Japanese case, for example, we assume that the anticipation of a Theme after NP2-dat entails that the event being described by this sentence is an act of transference. But does this mean that the processor projects a syntactic structure corresponding to the main verb? And that it projects a structure corresponding to NP3 that includes a specification of its case-marking? As just outlined, we cannot currently tell on any empirical basis. And as always, other suitable experimental paradigms are needed to spawn a body of converging evidence that may illuminate further the mechanisms by which the processor is sensitive to properties of the entities that might be referred to next.

Taken together, we believe that the most parsimonious account of our data is that the processor uses accruing constraints to compute the likely thematic relationships amongst the entities already referred to in the linguistic input, amongst the entities concurrently available in the visual context, and between the two. We view this process as a probabilistic one, which takes into account any available knowledge (linguistic or non-linguistic) concerning possible dependencies between the entities. In our Japanese examples, the interpretation of the sentential subject ('*waitress*-nom') as Agent, given its explicit case-marking, provides limited information regarding the dependencies amongst the remaining objects; the customer and/or the wastebin could be a Theme or Goal objects, the hamburger could be a Theme object, and so on. After NP2 ('*customer*-dat'), the processor is now much more constrained in respect of the possible thematic organization of the objects in the visual world. It can 'infer' (presumably on the basis of its experience of linguistic input) that a Theme object will be involved, and although it may be able to assume this just on the basis of encountering '*customer*-dat' (if you have a recipient, there must be something for the recipient to receive), it is more likely that it is assumed on the basis of the constraints afforded by the entire sequence 'NP-acc NP-dat'. Thus, by the time '*waitress*-nom' and '*customer*-dat' have been encountered, enough constraints based on their meaning and the grammatical cues they convey have been accrued to make the hamburger the most likely Theme object. Turning to the English case, we can describe a very similar pattern of constraint accrual—for a sequence such as *The girl will ride the carousel*, NP1 need not be an Agent (cf. *The girl was bought the* (*sweets*)), but even if it were interpreted as an Agent, the girl could plausibly engage in some activity with the man, the carousel, or the (sweets), and so NP1 alone would not usefully restrict the range of dependencies that may exist amongst the entities portrayed within the scene (or to be referred to in the subsequent linguistic input). However, the subsequent verb *does* restrict further the possibilities for subsequent reference, given its selectional restrictions—restrictions on the properties of the objects that can enter into a thematic relationship with, in this case, its Agent. Thus, constraints afforded by the verb (ride), coupled with those afforded by the Agent, lead to the carousel as the most likely object to enter into a dependency with that Agent. And although we cannot at present determine whether this entails a prediction that the carousel will in fact be referred to next (or whether in the Japanese case, the hamburger would be referred to next, with accusative case-marking), we believe nonetheless that computational demonstrations of constraint accrual, such as those exemplified by Elman (1990), suggest that such predictions can plausibly be made. Indeed, we view the processes that we have tapped in these studies as being similar (albeit on a different scale and across modalities) to the computational processes embodied in Elman's (1990) model. There, the system would predict what could come next as a function of the current input, and it did so using computational principles that are relatively transparent (see Altmann, 1997, both for an explanation of these principles and for an account of how the 1990 model could be extended to more adequately capture contextual dependencies within a multi-modal world).

Finally, and on the assumption that our findings generalize to situations other those with concurrent visual input, we believe that our findings reflect the nature of incrementality in human sentence processing. To interpret a sentence incrementally, word-by-word, requires the partial interpretation of what has been encountered thus far—thus, after *The boy will eat*, some representation must be constructed which at the very least places *eat* and *the boy* in a subject–predicate relationship. Beyond that, it may also reflect the semantic properties of *boy*, and of *eat* (including that the event being referred to takes place in the future). And given the semantics of *eat*, a partial interpretation may also reflect that fact that the boy will eat something. But the fullest possible interpretation would include also the interpretation at each moment in time of the consequences of what has been encountered so far for what may come next

(both in the language and in the mental or real world which that unfolding language describes). A full interpretation of the fragment *The boy will eat* must include the anticipation that something will be eaten, that it will be referred to post-verbally, that this post-verbal reference will take a particular form as licensed by the grammar, and that it is something that can plausibly be eaten by a boy (moreover, by this particular boy). Such an interpretation reflects, simply, all the syntactic, semantic, and real-world constraints that can be applied at the offset of this fragment. Although incrementality and prediction need not go hand in hand, we do nonetheless believe that the processes we have identified here and in our previous work are a distinguishing characteristic of an incremental processor that establishes the fullest possible interpretation at each moment in time.

Acknowledgments

We are grateful also to Vicki White, Lynne Weighall, and Matt Smelt-Webb for helping us recruit the participants in the experiments reported here. Also, we appreciate very helpful comments from Rick Lewis, and two anonymous reviewers given to earlier versions of the draft. The present research was supported by grants to GTMA from the Medical Research Council (G9628472N and G0000224), the Economic and Social Research Council (R000222798), and the Royal Society.

Appendix A

The experimental sentences in Experiment 1 are as follows. The *a* version of the sentences includes an Inanimate Goal, and the *b* version an Animate Goal. All the scenes used in the present paper (for Experiments 1–3) will be available on request from the first author.

1a. The artist will rest the picture on the easel.
1b. The artist will explain the picture to his wife.

2a. The boy will balance the lemons on the scales.
2b. The boy will pass the lemons to the woman.

3a. The man will chuck the can into the wastebin.
3b. The man will give the can to the cleaning lady.

4a. The woman will copy the message onto the floppy-disk.
4b. The woman will send the message to the man.

5a. The barman will drop the cloth onto the floor.
5b. The barman will toss the cloth to the customer.

6a. The older woman will hang the dress on the rack.
6b. The older woman will loan the dress to the teenage girl.

7a. The short-haired woman will lay the necklace in the jewellery case.
7b. The short-haired woman will sell the necklace to the other woman.

8a. The man will load the software onto the computer.
8b. The man will lend the software to the woman.

9a. The woman will lock the trumpet in the case.
9b. The woman will teach the trumpet to the boy.

10a. The woman will place the jumper in the chest of drawers.
10b. The woman will offer the jumper to the boy.

11a. The woman will position the sculpture on the table.
11b. The woman will return the sculpture to the man.

12a. The man will move the teddy bear onto the sofa.
12b. The man will present the teddy bear to his daughter.

13a. The man will arrange the books on the shelf.
13b. The man will read the books to his son.

14a. The woman will settle the baby into the pram.
14b. The woman will hand the baby to the boy.

15a. The older handyman will spray the paints onto the shelf.
15b. The older handyman will take the paints to his assistant.

16a. The woman will spread the butter on the bread.
16b. The woman will slide the butter to the man.

17a. The man will tack the poster on the wall.
17b. The man will show the poster to the woman.

18a. The waitress will throw the burger into the wastebin.
18b. The waitress will serve the burger to the man.

Appendix B

The 24 sets of experimental sentences used in Experiment 2 are as follows.

1a. The dog will bite the cat.
1b. The dog will drink the water.
1c. The man will bite the tomato.
1d. The man will drink the coffee.

2a. The girl will blow the windmill.
2b. The girl will ride the rocking horse.
2c. The man will blow the tuba.
2d. The man will ride the bicycle.

3a. The pirate will drink the brandy.
3b. The pirate will bury the treasure.
3c. The dog will drink the water.
3d. The dog will bury the bone.

4a. The puppy will chase the seagull
4b. The puppy will wear the collar.
4c. The Eskimo will chase the seal.
4d. The Eskimo will wear the gloves.

5a. The man will climb the mountains.
5b. The man will eat the orange.
5c. The cat will climb the tree.
5d. The cat will eat the mouse.

6a. The woman will close the door.
6b. The woman will drink the wine.
6c. The baby will close the jack-in-the-box.
6d. The baby will drink the milk.

7a. The cat will drink the water.
7b. The cat will climb up the fence.
7c. The builder will drink the tea.
7d. The builder will climb up the ladder.

8a. The boy will drive the toy car.
8b. The boy will fly the kite.
8c. The soldier will drive the tank.
8d. The soldier will fly the plane.

9a. The woman will eat the grapefruit.
9b. The woman will open the door.

9c. The cat will eat the bird.
9d. The cat will open the cat flap.

10a. The tourist will feed the child.
10b. The tourist will close the suitcase.
10c. The zoo keeper will feed the elephant.
10d. The zoo keeper will close the gate.

11a. The paramedic will wear the stethoscope.
11b. The paramedic will help the patient.
11c. The fireman will wear the helmet.
11d. The fireman will help the car owner.

12a. The dolphin will jump through the sea.
12b. The dolphin will eat the fish.
12c. The monkey will jump through the trees.
12d. The monkey will eat the bananas.

13a. The bird will drink the water.
13b. The bird will line the nest.
13c. The woman will drink the tea.
13d. The woman will line the jacket.

14a. The little girl will open the present.
14b. The little girl will dress the doll.
14c. The adult woman will open the champagne.
14d. The adult woman will dress the baby.

15a. The worker will read the plans.
15b. The worker will build the house.
15c. The boy will read the comic.
15d. The boy will build the sand castle.

16a. The tailor will repair the sawing machine.
16b. The tailor will cut the fabric.
16c. The plumber will repair the sink.
16d. The plumber will cut the pipe.

17a. The farmer will ride the tractor.
17b. The farmer will bite the carrots.
17c. The boy will ride the unicycle.
17d. The boy will bite the ice lolly.

18a. The captain will sail the boat.
18b. The captain will wear the gloves.
18c. The boy will sail the toy yacht.
18d. The boy will wear the cap.

19a. The boy will shoot the archery target.
19b. The boy will sip the coke.
19c. The hunter will shoot the fox.
19d. The hunter will sip the gin.

20a. The boy will sleep in the bed.
20b. The boy will devour the water melon.
20c. The dog will sleep in the kennel.
20d. The dog will devour the dog food.

21a. The bird will swallow the worm.
21b. The bird will feed the chicks.
21c. The woman will swallow the vitamins.
21d. The woman will feed the baby.

22a. The frog will eat the fly.
22b. The frog will swim in the pond.
22c. The girl will eat the ice cream.
22d. The girl will swim in the pool.

23a. The man will taste the beer.
23b. The man will ride the motorbike.
23c. The girl will taste the sweets.
23d. The girl will ride the carousel.

24a. The man will wear the boxer shorts.
24b. The man will read the car magazine.
24c. The woman will wear the dress.
24d. The woman will read the fashion magazine.

Appendix C

The experimental sentences in Experiment 3 are as follows. The *a* version of the sentences is the Dative condition, and the *b* version is the Accusative condition.

1a. ウェイトレスが客に楽し気にハンバーガーを運ぶ。
weitoresu-ga kyaku-ni tanosigeni hanbaagaa-o hakobu.
waitress-nom customer-dat merrily hamburger-acc bring.
'The waitress will merrily bring the hamburger to the customer.'

1b. ウェイトレスが客を楽し気にからかう。
weitoresu-ga kyaku-o tanosigeni karakau.
waitress-nom customer-acc merrily tease.
'The waitress will merrily tease the customer.'

2a. 男性が壁にていねいにポスターを貼る。
dansee-ga kabe-ni teeneeni posutaa-o haru.
man-nom wall-dat carefully poster-acc hung.
'The man will carefully hung the poster on the wall.'

2b. 男性が壁をていねいに拭く。
dansee-ga kabe-o teeneeni huku.
man-nom wall-acc carefully wipe.
'The man will carefully wipe the wall.'

3a. 少女が庭にゆっくりと椅子を運び出す。
syoojo-ga niwa-ni yukkurito isu-o hakobi-dasu.
girl-nom garden-dat slowly chair-acc move out.
'The girl will slowly move the chair out to the garden.'

3b. 少女が庭をゆっくりと眺める。
syoojo-ga niwa-o yukkurito nagameru.
girl-nom garden-acc slowly look at.
'The girl will slowly look at the garden.'

4a. 女性がたんすにてきぱきとセーターを入れる。
josee-ga tansu-ni tekipakito seetaa-o ireru.
woman-nom chest of drawers-dat efficiently sweater-acc put.
'The woman will efficiently put the sweater into the chest of drawers.'

4b. 女性がたんすをてきぱきと移動する。
josee-ga tansu-o tekipakito idoo-suru.
woman-nom chest of drawers-acc efficiently move.
'The woman will efficiently move the chest of drawers.'

5a. 男性が椅子に大雑把にスプレーをかける。
dansee-ga isu-ni oozappani supuree-o kakeru.
man-nom chair-dat roughly spray-acc put.
'The man will roughly spray the arsenal on the chair.'

5b. 男性が椅子を大雑把に蹴飛ばす。
dansee-ga isu-o oozappani ke-tobasu.
man-nom chair-acc roughly kick away.
'The man will roughly kick the chair away.'

6a. 未亡人が男性に喜んで置き物を貸す。
miboozin-ga dansee-ni yorokonde okimono-o kasu.
widow-nom man-dat willingly ornament-acc lend.
'The widow will willingly lend the ornament to the man.'

6b. 未亡人が男性を喜んで招待する。
miboozin-ga dansee-o yorokonde syootai-suru.
widow-nom man-acc willingly invite.
'The widow will willingly invite the man.'

7a. 女性が少年にうれしそうにハープを見せる。
josee-ga syoonen-ni uresisooni haapu-o miseru.
woman-nom boy-dat delightedly harp-acc show.
'The woman will delightedly show the harp to the boy.'
7b. 女性が少年をうれしそうに歓迎する。
josee-ga syoonen-o uresisooni kangee-suru.
woman-nom boy-acc delightedly welcome.
'The woman will delightedly welcome the boy.'

8a. 男性が妹にゆっくりと椅子を持っていく。
dansee-ga imooto-ni yukkurito isu-o motte-iku.
man-nom sister-dat slowly chair-acc bring.
'The man will slowly bring the chair to his sister.'
8b. 男性が妹をゆっくりと撫でる。
dansee-ga imooto-o yukkurito naderu.
man-nom sister-acc slowly stroke.
'The man will slowly stroke his sister.'

9a. バーテンダーが客にいきなりふきんを投げる。
baatendaa-ga kyaku-ni ikinari hukin-o nageru.
bartender-nom customer-dat abruptly cloth-acc throw.
'The bartender will abruptly throw the cloth at the customer.'
9b. バーテンダーが客をいきなり追い返す。
baatendaa-ga kyaku-o ikinari oi-kaesu.
bartender-nom customer-acc abruptly kick out.
'The bartender will abruptly kick out the customer.'

10a. 芸術家がテーブルに何気なく絵を置く。
geejutuka-ga teeburu-ni nanigenaku e-o oku.
artist-nom table-dat unconcernedly painting-acc put.
'The artist will unconcernedly put the painting onto the table.'
10b. 芸術家がテーブルを何気なく見つめる。
geejutuka-ga teeburu-o nanigenaku mitumeru.
artist-nom table-acc unconcernedly stare at.
'The artist will unconcernedly stare at the table.'

11a. 女性が男性にこっそりとメッセージを送る。
josee-ga dansee-ni kossorito messezi-o okuru.
woman-nom man-dat secretly message-acc send.
'The woman will secretly send the message to the man.'
11b. 女性が男性をこっそりと好きになる。
josee-ga dansee-o kossorito sukini-naru.
woman-nom man-acc secretly fall for.
'The woman will secretly fall for the man.'

12a. 男性がガレージに注意深くビールを収納する。
dansee-ga gareezi-ni chuuibukaku biiru-o syuunoo-suru.
man-nom garage-dat carefully beer-acc store.
'The man will carefully store the beer into the garage.'
12b. 男性がガレージを注意深く点検する。
dansee-ga gareezi-o chuuibukaku tenken-suru.
man-nom garage-acc carefully check.
'The man will carefully check the garage.'

13a. 男性社員がゴミ箱に乱暴に空き缶を捨てる。
dansee syain-ga gomibako-ni ranbooni akikan-o suteru.
male employee-nom dust bin-dat violently empty can-acc throw away.
'The male employee will violently throw away the empty can into the dust bin.'
13b. 男性社員がゴミ箱を乱暴に蹴り倒す。
dansee syain-ga gomibako-o ranbooni keri-taosu.
male employee-nom dust bin-acc violently kick down.
'The male employee will violently kick the dust bin down.'

14a. 男性がコンピューターにようやくソフトウェアを立ち上げる。
dansee-ga konpuutaa-ni yooyaku sohutowea-o tati-ageru.
man-nom computer-dat finally software load up.
'The man will finally load up the software on the computer.'
14b. 男性がコンピューターをようやく使い始める。
dansee-ga konpuutaa-o yooyaku tukai-hazimeru.
man-nom computer-acc finally use start.
'The man will finally start to use the computer.'

15a. 女性が男性にうれしそうに論文を渡す。
josee-ga dansee-ni uresisoo-ni ronbun-o watasu.
woman-nom man-dat delightedly thesis-acc pass.
'The woman will delightedly pass the thesis on to the man.'
15b. 女性が男性をうれしそうに見つめる。
josee-ga dansee-o uresisoo-ni mi-tumeru.
woman-nom man-acc delightedly look into.
'The woman will delightedly look into the man.'

16a. 年配女性が若い女性にいやいやドレスを譲る。
nenpai josee-ga wakai josee-ni iyaiya doresu-o yuzuru.
elderly woman-nom young woman-dat unwillingly dress-acc hand over.
'The elderly woman will unwillingly hand over the dress to the young woman.'
16b. 年配女性が若い女性をいやいやほめる。
nenpai josee-ga wakai josee-o iyaiya homeru.
elderly woman-nom young woman-acc unwillingly praise.
'The elderly woman will unwillingly praise the young woman.'

References

Altmann, G. T. M., Garnham, A., & Henstra, J. A. (1994). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*(1), 209–216.

Altmann, G. T. M. (1997). *The ascent of Babel*. Oxford: Oxford University Press.

Altmann, G. T. M. (1999). Thematic role assignment in context. *Journal of Memory and Language, 41*, 124–145.

Altmann, G. T. M., Garnham, A., van Nice, K., & Henstra, J. A. (1998). Late closure in context. *Journal of Memory and Language, 38*(4), 459–484.

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.

Androutsopoulos, I., & Dale, R. (2000). Selectional restrictions in HPSG'. Proceedings of the 18th International Conference on Computational Linguistics (COLING-2000) (pp. 15–20).

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

Bertolucci, B. (1972). Last Tango in Paris. United Artists.

Boland, J. E., Tanenhaus, M. K., & Garnsey, S. M. (1990). Verb argument structure in parsing and interpretation. Evidence from wh-questions. *Journal of Memory and Language, 34*, 774–806.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language, 34*, 774–806.

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.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: The MIT Press.

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.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211.

Hemforth, B. (1993). *Kognitives parsing: Repräsentation und verarbeitung sprachlichen wissens*. Sankt Augustin: Infix.

Jackendoff, R. S. (1977). *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Kamide, Y., & Altmann, G. T. M., (2000, September) Predictive projection of subsequent constituents: Evidence from anticipatory eye movements during Japanese parsing. Poster presented at the Architectures and Mechanisms for Language Processing (AMLaP)-2000 Conference, Leiden, the Netherlands.

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

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32*, 37–55.

Kamide, Y., Altmann, G. T .M. (in preparation). Argument structure biases in English ditransitive constructions: Norms and their application to sentence processing research.

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

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676–703.

Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: information processing time with and without saccades. *Perception & Psychophysics, 53*, 372–380.

Mazuka, R., Itoh, K., & Kondo, T. (2000). Costs of scrambling in Japanese sentence processing. In M. Nakayama (Ed.), *Japanese/Korean Linguistics* (Vol. 9 (9),. London: CSLI Publications.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 134–157.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes, 12*, 137–176.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and performance* (Vol. 12,. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mitchell, D. C., Corley, M., & Garnham, A. (1992). Effects of context in human sentence processing: Evidence against a discourse-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 69–88.

Miyamoto, E. T., & Takahashi, S. (2002). Sources of difficulty in processing scrambling in Japanese. In M. Nakayama (Ed.), *Japanese/Korean Linguistics* (Vol. 9 (9),. London: CSLI Publications.

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago, London: University of Chicago Press and CSLI Publications.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 358–374.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America, 57*(8), 1030–1033.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*(2), 109–148.

Steedman, M. J. (2000). *The syntactic process*. Cambridge, MA: MIT Press/Bradford Books.

Tanenhaus, M. K., Carlson, G., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes, 4*, 211–234.

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.

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 and Cognition, 19*, 528–553.

Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behaviour, 16*, 683–692.