Conflicting Constraints in Resource-Adaptive Language Comprehension

Andrea Weber, Matthew W. Crocker, and Pia Knoeferle

1 Introduction

The primary goal of psycholinguistic research is to understand the architectures and mechanisms that underlie human language comprehension and production. This entails an understanding of how linguistic knowledge is represented and organized in the brain and a theory of how that knowledge is accessed when we use language. Research has traditionally emphasized purely linguistic aspects of on-line comprehension, such as the influence of lexical, syntactic, semantic and discourse constraints, and their time-course. It has become increasingly clear, however, that non-linguistic information, such as the visual environment, are also actively exploited by situated language comprehenders. The wealth of informational resources which are potentially relevant to situated language comprehension raise a number of important questions. To what extent are the mechanisms underlying comprehension able to exploit linguistic and non-linguistic information on-line, and how do people adapt to the availability or non-availability of contextual information?

We begin below, with a brief summary of several important aspects of human language comprehension, including its incremental and even anticipatory nature, and its sensitivity to accrued linguistic experience. We then present a range of experimental findings which reveal the ability of comprehenders to rapidly adapt to diverse linguistic and non-linguistic constraints. To better understand this apparently seamless ability of the human language processing faculty to integrate diverse cues, including linguistic context, intonation, world knowledge, and visual context, many of the experiments are designed so as to better understand the relative priority of these constraints when they are pitted against each other. The findings conspire to paint a picture in which purely linguistic constraints, long thought to identify the core of sentence comprehension mechanisms, can in fact be overridden by highly contextual aspects of the situation, such as the intonation contour of a particular utterance, semantic expectations supported by the visual scene, and indeed events going on in the scene itself.

A. Weber (✉)
Max-Planck-Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands
e-mail: andrea.weber@mpi.nl

M.W. Crocker, J. Siekmann (eds.), *Resource-Adaptive Cognitive Processes,*
Cognitive Technologies, DOI 10.1007/978-3-540-89408-7_7,

1.1 Incrementality

A basic finding is that human sentence processing is incremental. That is, humans structure and interpret the words of an utterance as they are perceived rather than store them as a list to be combined later. A seminal finding for incrementality was published in 1973 by Marslen-Wilson [24], who showed in a speech-shadowing experiment that both syntactic and semantic information are available to participants as they repeat the speech they hear; constructive errors were usually grammatically suitable with respect to the preceding context, even for shadowers who repeated the speech input with a minimal time-lag. This suggests that the shadowers' performance was based on a syntactic analysis of the ongoing speech stream. Since that time, empirical support for the claim that language comprehension takes place incrementally is overwhelming. Evidence from eye-tracking has shown that people not only rapidly map the unfolding words onto visually present objects, but that they also structure the words of an utterance into a connected interpretation as they are encountered (e.g., [35]), and they even have expectations about the words they predict to come (e.g., [2]).

However, while incremental processing and interpretation ensures real-time understanding, it brings with it additional challenges. Sequences are often ambiguous; that is, they are compatible with more than one well-formed structural representation. For example, in the sentence beginning *Betty knew Monica's date* ..., *Monica's date* could be the direct object of *knew* or could become the subject of a clausal complement (*Betty knew Monica's date had bought flowers*). Disambiguating information may occur in later parts of the sentence, but due to incrementality, processing must proceed before such relevant information becomes available. A great deal of research has therefore focused on the processing of local ambiguities as a means for investigating the kinds of information and strategies listeners employ during the earliest stages of sentence processing [13, 7, 27].

1.2 Multiple Constraints

An abundance of empirical studies have specified the different information sources that are used on-line for ambiguity resolution in sentence processing. For example, it has been shown numerous times that the human parser resolves structural ambiguities using a set of processing preferences. One of these is a preference to always build the simplest structure; a direct object attachment of a postverbal noun phrase (NP) is, for example, less complex than an attachment that would require the additional structure associated with a complement clause (the parser thus prefers to analyze *Monica's date* as direct object in *Betty knew Monica's date*). This preference is known as the Minimal Attachment principle [13].

Besides purely structural information, prior linguistic experience has been shown to play an important role in human sentence processing. There is wide-ranging evidence, for example, that frequency-derived lexical preferences influence the processing of ambiguity. Reconsider the *Betty knew Monica's date* example, and replace the verb with *thought*. *Think* is a verb that cannot be followed by a direct object, but only by a clausal complement. Thus *Betty thought Monica's date* would be given a clausal complement analysis, because only a clause such as *had bought flowers* can follow; no ambiguity would be encountered. When verbs can appear in more than one structure (e.g., *admit* can either appear with a direct object or with a clausal complement), empirical research has shown that the structural processor can be biased in its initial analysis towards the more frequent sentences type for this verb (e.g., [14]; but for contradictory results see [30]). More generally, Crocker argues that the pervasive role of experience, as supported by a range of findings revealing the influence of frequency, offers a fundamental explanation for how people adapt to language over time, enabling them to deal so effectively with ambiguity [8]. Both probabilistic [9, 6, 8] and connectionist models of sentence processing [11] (see also Mayberry and Crocker, *The Evolution of a Connectionist Model of Situated Human Language Understanding* of this volume) naturally manifest the central role of experience, favoring interpretations which are supported by evidence they were exposed to during training. As a consequence, experience-based modes fit with a rational view of linguistic performance in which processing mechanisms seek to optimize understanding [5], by recovery of the interpretation most likely to correct, rather than minimize representational or processing complexity [13, 15].

In addition, conceptual knowledge has been shown to influence the processing of syntactic ambiguity. In particular the assignment of thematic roles to noun phrases has served as a test case (e.g., [31, 27]). The thematic roles of a verb describe the mode of participation entities play in the event denoted by the verb: For example, cops usually arrest criminals (and are therefore suitable agents for the event of arresting) whereas criminals usually are being arrested (and are therefore suitable patients for the event of arresting). Reading times in [27], for example, suggest that readers compute and use such event-specific world knowledge immediately for the interpretation of the ambiguous region of reduced relative clauses (*The criminal/cop arrested by* ...) as evidenced by modulation of reading times in the subsequent disambiguation region.

A fourth important factor for resolving structural ambiguity is discourse context. For example, the sentence *Monica told her friend that she had been trying to avoid* ... could be completed with *her date* or with *to call back tomorrow*. In the first case *that she had been trying to avoid* would be an assertion told to Monica's friend; in the later case the same phrase would be a specification for which friend Monica meant. Altmann et al. [1] found that discourse context plays an important part in determining how these sentences are read. For example, if Monica had previously mentioned two friends, then *that she had been trying to avoid* is analyzed by listeners as a distinguishing modification. These findings have also been replicated in situated spoken language comprehension, where the relevant referential context is provided by a visual scene, rather than a prior discourse, crucially highlighting comprehenders' ability to exploit both linguistic and non-linguistic context [35].

This partial survey of psycholinguistic findings, clearly support the notion of a human sentence processing mechanism that is not only incremental but is also highly adaptive to different information sources. Both constraints resulting from long-term exposure to language, like biases for lexical verb frames or preferences for certain syntactic structures, as well as constraints resulting from short-term exposure, like discourse context, are rapidly exploited as they become available during on-line sentence processing.

1.3 Anticipation in Situated Comprehension

More recently, evidence is mounting that sentence processing is not only incremental, but even anticipatory (e.g., [2]): Listeners are able to rapidly combine information from different constituents to predict subsequent arguments. While the more traditional experimental method of tracking eye movements during reading provided detailed information about the time course of various interpretation processes, anticipatory behavior was not easily detectable with this method. With the advent of eye-tracking in visual scenes [35], however, it became possible to gain clear insight into both the current interpretation listeners are adopting, as well as continuations of a sentence that they expect would plausibly follow from that interpretation. Whereas in reading studies, text is displayed on a computer screen and reading times at different positions in the text (usually the point of disambiguation) allow conclusions about cognitive processing load, in *visual-world* studies participants view scenes depicting objects and events while simultaneously listening to a related utterance. Eye movements are measured in relation to interesting regions of the acoustically presented sentence, such as a noun referring to the objects on the screen. Such utterane-mediated gaze is closely time-locked with the unfolding sentence, with shifts in visual attention occurring about 200 ms after the relevant spoken material is heard. Empirical evidence has further shown that listeners make eye movements in anticipation that a picture in a display will become relevant. For example, upon hearing *the boy will eat*, listeners start looking at edible objects even before they are mentioned [2]. Anticipatory eye movements can thus inform us about higher-level processes, such as the role of verb information in restricting the domain of subsequent reference.

2 Varying Constraints

Outside the laboratory, in the real world, language users have to deal with multiple information sources and modalities simultaneously. Everyday sentences include structural, lexical, discourse, as well as prosodic information in varying degrees; the listeners' task is then to successfully use the relevant information to guide sentence interpretation. It is likely that the impact of different information types changes with varying circumstances; also one information type might be more important than another type, and their impact might happen at different times in the sentence.

Ultimately, any theory of human sentence processing must be able to account for sentence processing in the light of multiple, varying information sources. In responding to this, psycholinguistic research therefore needs to shift away from simply establishing which information sources influence on-line sentence processing, and place increased emphasis on determining the circumstances under which each type of information source is more or less likely to have an impact. One approach that will bring us closer to achieving this goal is to study comprehension in the face of varying and even contradictory information sources. In this way we explore the extent to which specific information types are favored, dismissed, or weighted with respect to each other. A secondary issue concerns the notion of *task*, as evidence mounts for the view that people process language in importantly different ways depending on whether they are simply reading [32], required to make judgements or answer questions [34], or even to carry out spoken instructions [35]. The exploration and development of such an account of adaptive mechanisms in sentence processing will thus better account for variations in behavior in diverse contexts and tasks.

We present five representative experimental investigations conducted in the context of the ALPHA project that address the issue of sentence processing in light of varying information sources. The first study was concerned with the role of discourse information in word ordering preferences. In this project, we tested whether difficulties with processing non-canonical word orders in German can be weakened with discourse context which provides information about grammatical functions. The second study investigated the interaction of syntactic ordering preferences with prosodic information: In spoken language, intonation contours can convey a range of communicative functions. We tested whether listeners rely on a specific prosodic pattern for the interpretation of German scrambled sentences. The third study looked at the influence of lexical preferences on semantically constrained verb arguments. Semantic verb information is known to restrict listeners' expectations about upcoming verb arguments, and we examined in this study the role of experience with lexical items in forming argument expectations. In the fourth study, the influence of scene objects on linguistic expectations was examined. Whereas in the third study we assessed the long-term constraint of lexical frequency in semantically constraining contexts, in the fourth study we tested the short-term constraint of visual context in semantically constraining utterances. Finally, in the fifth set of experiments, we more deeply investigate the on-line interplay of scene and language processing, and examine the priority of scene information relative to expectations arising from our longer-term world knowledge.

2.1 Discourse Information and Structural Preferences

German is a language with relatively free constituent order. For instance, the initial position in matrix declaratives observes very few restrictions regarding the kind of constituent it can host, which includes subjects, objects, as well as modifiers. Thus both SVO orders like *der Verein St. Johann gewann den Pokal*, "the club$_{NOM}$ St. Johann won the prize$_{ACC}$" and OVS orders like *den Pokal gewann der Verein St. Johann*, "the prize$_{ACC}$ won the club St. Johann$_{NOM}$" are possible in German, though there is clear preference for the canonical subject-first order (see, e.g., [17]). In the previous example, the nominative case of the subject as well as the accusative case of the object are unambiguously assigned with case marking. However, although German does use morphological case to mark grammatical functions, the system often features syncretism: In many noun phrases (NPs), nominative and accusative cases share surface form. As a result, the constituent ordering of a sentence can be ambiguous: *die Mutter ruft die Tochter*, "the mother$_{NOM,ACC}$ calls the daughter$_{NOM,ACC}$" could either mean that the mother is calling the daughter (SVO) or that the daughter is calling the mother (OVS). In order to correctly interpret an utterance in which the structure cannot be determined on the basis of linguistic information alone, human language users may rely on other information sources to resolve the ambiguity. One such short-term information source might be discourse context. Weber and Neu [40] tested this assumption in a German reading study in which information about grammatical functions of referents could only be inferred from information in the preceding discourse.

In their study, a target sentence with a temporal word order ambiguity was preceded by a question that assigned the grammatical function to one of the referents in the target sentence. In the target sentences, case marking of initial NPs was ambiguous with respect to grammatical function, while case marking of the second NP disambiguated sentences towards SO or OS order (e.g., *Die Katze jagt gleich den Vogel/der Hund mit grossem Eifer*, "the cat$_{NOM,ACC}$ chases in-a-moment the bird$_{ACC}$/the dog$_{NOM}$ with great eagerness"). Without further context, the default interpretation of *die Katze* is subject, since subject-first sentences are the canonical order in German; no processing difficulties should arise upon reading the second object argument *den Vogel*, since it agrees with the subject interpretation of *die Katze*. However, upon encountering a subject as second argument (*der Hund*), readers will have to revise their initial interpretation of *die Katze* as subject; this will be reflected in longer reading times of the second argument *der Hund*. Preceding context consisted of two sentences: a declarative sentence introducing three possible referents (e.g., *Auf der Wiese sind eine Katze, ein Hund und ein Vogel*, "on the field are a cat, a dog and a bird"), followed by a focussing wh-question. Crucially, the focussing question provided information about the grammatical function of a subsequent referent. For instance, the question *Wen jagt gleich die Katze mit grossem Eifer?*, "whom$_{ACC}$ chases in-a-moment the cat with great eagerness?" introduces the cat as subject, the grammatical role the cat will most likely also take in a subsequent answer. The question particles of the focussing questions were either *who* (NOM) or *whom* (ACC). In a baseline condition, a question that did not assign grammatical functions to subsequent NPs was used. For an example of a complete stimulus set see Fig. 1.

There were two reasons for using different question types: For one, both the *who* and the *whom* questions were providing information about the grammatical function of the first NP in the target sentences, whereas the baseline questions did not provide such information. A comparison of focussing questions with the baseline question would therefore inform us whether the processing of sentences with canonical and non-canonical words orders profits from contextual focus. The comparison between *who* and *whom* questions, on the other hand, would inform us about the additional

wh_NOM	Wer jagt gleich den Vogel mit großem Eifer? *Who (NOM) chases in-a-moment the bird with great eagerness?*
wh_ACC	Wen jagt gleich die Katze mit großem Eifer? *Whom (ACC) chases in-a-moment the cat with great eagerness?*
wh_neutr	Was passiert gleich? *What will in-a-moment happen?*
target SO	**Die Katze jagt gleich den Vogel mit großem Eifer.** *The cat (NOM, ambiguous) chases in-a-moment the bird (ACC) with great eagerness.*
wh_NOM	Wer jagt gleich die Katze mit großem Eifer? *Who (NOM) chases in-a-moment the cat with great eagerness?*
wh_ACC	Wen jagt gleich der Hund mit großem Eifer? *Whom (ACC) chases in-a-moment the dog with great eagerness?*
wh_neutr	Was passiert gleich? *What will in-a-moment happen?*
target OS	**Die Katze jagt gleich der Hund mit großem Eifer.** *The cat (ACC, ambiguous) chases in-a-moment the dog (NOM) with great eagernes.*

Fig. 1 Example of stimulus set with three different questions preceding both the SO target sentence and the OS target sentence

influence of structural expectancies; whereas after *who* questions answers are more likely to begin with the subject (SO), after *whom* questions, the object is more likely to be in sentence-initial positions (OS).

Weber and Neu [40] found faster total reading times for the second NP in target sentences (the point of disambiguation) when sentences were preceded by focussing questions (*who* or *whom*) than by the baseline question (see Fig. 2). This supports the assumption that both locally ambiguous canonical and non-canonical word orders

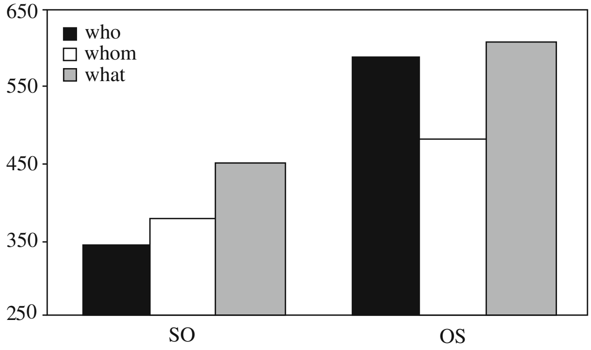

Fig. 2 Total reading times in ms for the second disambiguating NP in SO and OS sentences after a focusing who- or whom-question, and a neutral what-question

profit from focus in a preceding discourse context. Note, however, that OS sentences preceded by focussing questions were still harder to process than comparable SO sentences. So the difficulties of processing non-canonical word orders were not fully overcome by focussing context.

Second, both SO and OS sentences were easier to process when the syntactic structure of the focussing question was matching with the structure of the target sentence. For SO sentences, reading times were faster when the sentence was preceded by a *who* question than by a *whom* question; for OS sentences, reading times were faster when the sentence was preceded by a *whom* question than by a *who* question. This result is in line with findings of syntactic priming in comprehension in which sentences were found to be processed more easily when they were preceded by sentences with a matching syntactic structure than by a mismatching syntactic structure (e.g., [4]). However, when syntactic structure was mismatching, SO sentences in [40] were still easier to process than the baseline condition. Thus, processing can still gain from information about grammatical functions even when there is a structural mismatch. And finally a ray of hope for the processing of non-canonical OS sentences: even though OS sentences were overall more difficult to comprehend than SO sentences, the presence of a focusing question that also matched in syntactic structure resulted at least in reading times that were comparable with the baseline condition of the canonical SO sentences. Thus, short-term information from discourse context can significantly help to overcome processing difficulties with non-canonical word orders in German.

2.2 Prosodic Information and Structural Preferences

A further short-term information source in spoken sentence processing, besides discourse information, is prosody. Prosody is the description of phrasing, stress, loudness, and the placement and nature of pitch accents in spoken language. It can express or aid a range of functions in communication: mark the difference between immediately relevant vs. background information, express contrast, contradiction, correction, or even indicate the intended syntax of ambiguous utterances. Prosody is different from the other information sources in that it is highly variable in its realization. There is, for instance, no simple and direct correspondence between syntactic and prosodic structures. Quite often, a speaker can choose between a number of different intonation contours to express a particular communicative function. Nevertheless, it has been shown that listeners rely on prosodic information in sentence processing. On a structural level, for example, evidence has been presented that prosody can guide listeners' interpretation of attachment ambiguities (e.g., [19]). Sentences with early closure (*When Roger leaves the house is dark*) were compared with late closure sentences (*When Roger leaves the house it's dark*), and using a variety of experimental tasks it was shown that sentences with cooperating prosody (i.e., with a prosodic boundary after *leaves* in the early closure sentence) were processed more quickly than those with baseline prosody. Sentences with conflicting prosody were processed more slowly than those with baseline prosody.

Weber et al. [38] examined the role of prosody in a different ambiguity type, namely word order ambiguity in German. Incorrect initial interpretation of word order typically results in a much stronger garden-path effect than the previously tested modifier attachment ambiguities. One possible reason for this is that reanalysis from an SVO to an OVS structure entails a complete reassignment of the verbs' roles to both arguments. Given the stronger-garden path effect, it is particularly interesting to attest the role of prosody in this ambiguity type.

As described in the previous section, German nominative and accusative case often share surface forms. In combination with free constituent order in German, a functional gap arises: for example, *die Katze*, "the cat", in utterance initial position can be both subject (nominative case) and object (accusative case). In an eye-tracking study with visual scenes, Weber et al. [38] examined whether prosody can fill the functional gap arising from a combination of syncretism and free constituent order in German. Can prosody, in the absence of unambiguous morphological and configurational information, influence the assignment of grammatical function?

To investigate this question, they observed anticipatory eye movements of German listeners in a scene during comprehension of a related utterance. Not only has it repeatedly been shown that referents in a scene are identified as soon as they are referred to in an utterance, there are several studies revealing that they can be identified prior to their mention. With respect to constituent order ambiguity in German, two eye-tracking studies priorly attested such anticipatory behavior. For one, Kamide et al. [18] have shown that unambiguous case marking, combined with verb selectional information, leads to post-verbal anticipatory eye movements in German SVO and OVS sentences. That is, upon hearing *der Hase frisst...*, "the hare$_{NOM}$ eats...", German participants start to look at an appropriate object argument in the scene (e.g., a cabbage) even before hearing the second argument; upon hearing *den Hasen frisst...*, "the hare$_{ACC}$ eats...", they anticipate an appropriate subject argument (e.g., a fox). Thus, listeners are able to use case marking to assign the appropriate grammatical function to the first acrgument and combine this with the semantics of the verb, resulting in increased anticipatory fixations to the appropriate second argument.

Weber et al. [38] similarly employed German SVO and OVS structures, but with sentence-initial NPs that were ambiguously marked for nominative or accusative case. Morphosyntactic disambiguation of grammatical functions took place at the second NP that was clearly case marked as either accusative or nominative (e.g., *Die Katze jagt womöglich den Vogel/der Hund*, "the cat$_{NOM,ACC}$ chases possibly the bird$_{ACC}$/the dog$_{NOM}$"). Scenes accompanying the sentences showed the referent of the first NP (e.g., a cat) and plausible objects and subjects for the referent of the first NP in relation to a given action (e.g., a bird as plausible object for being chased by a cat and a dog as plausible subject for chasing a cat, see Fig. 3). No actions were depicted. Thus, even though the scenes presented potential referents they could not help with disambiguating grammatical roles in any way (see Sect. 2.5). In contrast with previous studies, however, prosodic cues could potentially help listeners resolve the temporary SVO/OVS ambiguity.

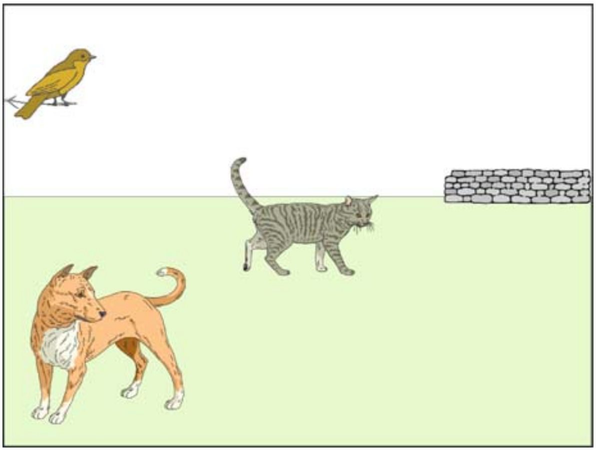

Fig. 3 Visual context for spoken sentences: *Die Katze jagt womöglich den Vogel/derHund*

The SVO sentences had a low pitch accent on the first NP (L* + H according to GToBI transcription [16]), followed by a focal high-pitch accent (H*) on the verb. This prosodic pattern was considered unmarked and was expected to indicate canonical subject-first sentences. The OVS sentences had a focal high-pitch accent (L + H*) on the first NP. This prosodic pattern was considered marked and was expected to indicate non-canonical object-first sentences. During the verb (e.g., chases), no effect of prosody was found. That is, potential objects (e.g., a bird) were fixated more often than potential agents (e.g., a dog) in both SVO and OVS sentences. Looks to the potential object imply that the initial NP was interpreted as subject (and therefore agent). The preference for anticipatory looks to potential objects at this point is a mere reflection for the well-attested preference for the canonical SVO order in German. During the following adverb (e.g., possibly), however, only for SVO structure more looks to potential objects were found. For OVS structures, potential subjects drew slightly more looks than potential objects. Thus, the strong preference for an SVO interpretation disappeared in sentences with OVS-type intonation. Prosodic cues were interpreted rapidly enough to affect listeners' interpretation of grammatical function before disambiguating case information was available.

The influence of prosodic information on the resolution of word order ambiguity is particularly striking for two reasons. First, the preference for the canonical SVO structure is very strong for German listeners. This is not surprising given that only about 18% of German sentences are OVS (in the Negra corpus; [39]). Most likely, this preference is stronger than that of previously tested attachment ambiguities. Prosodic information is therefore competing against a structural preference that has found plenty of support from long-term exposure to language and is highly ingrained. Second, as mentioned before, prosodic realizations are variable. A nuclear pitch accent on the first NP is definitely not the only way to intone an OVS structure. Intonation contours with phrase breaks or silent intervals after the first NP are also easily imaginable, for example. In addition, a nuclear pitch accent on the first NP can have in a different context a different meaning; for instance, the same pitch accent is known to convey contrasts. Further research is necessary to test whether other prosodic patterns can similarly influence the interpretation of grammatical functions. For the specific situation of the described study, however, we could show that prosodic focus on the first NP in OVS sentences placed a high prominence on the noun phrase which in turn facilitated the interpretation of the marked syntactic structure OVS.

2.3 Semantic Information and Lexical Preferences

Similar to the anticipation of arguments based on grammatical information we described above, anticipatory behavior in eye-tracking studies has been found for semantically constraining verb information. That is, listeners start looking at pictures of suitable object NPs right after semantically constraining verbs [2]: following *the boy will eat*, listeners fixate edible objects in a scene even before they are mentioned in the utterance. The semantic information extracted at the verb is sufficient to exclude other visually presented objects as potential referents. This entails that the human processor can immediately establish anaphoric dependencies on the basis of thematic fit between the referents in the visual context and the verb.

At the same time, there is ample evidence that the human processor has lexical biases which are built on long-term experience; words that occur more often in a language are favored and recognized more easily than less frequent words (e.g., [25]). In particular, the simultaneous activation of word candidates with overlapping onset has been shown to be modulated by lexical frequency [10]: While hearing the word *bench*, English listeners look more at the distractor picture of the high-frequency *bed* than at the distractor picture of the low-frequency *bell* in an eye-tracking study. The combination of semantic information and lexical preferences can lead to a situation in which verb information constrains potential referents in the presence of semantically inapt high-frequency distractors. Weber and Crocker [36] investigated the interaction of lexical frequency effects with effects from verb constraints in a German eye-tracking study with visual scenes. In particular, they tested whether high-frequency distractors are activated even though semantic information from preceding verbs renders them unlikely word candidates.

In their study, German participants listened to sentences with restrictive and unrestrictive verbs (e.g., *Die Frau bügelt/sieht die Bluse*, “the woman is ironing/seeing the blouse”) while they were looking at a display with four objects. The display showed the agent of the sentence (e.g., *Frau*, “woman”), a low frequency target (e.g., *Bluse*, “blouse”), a high-frequency phonological distractor (e.g., *Blume*, “flower”), and an unrelated distractor (e.g., *Wolke*, “cloud”; high in lexical frequency but phonologically unrelated to the target) (see Fig. 4). From the view of semantic information, the target and the distractors are possible object arguments following the unrestrictive verb (e.g., is seeing), but only the target is a likely candidate following the restrictive verb (e.g., is ironing). From the view of lexical frequency, however,

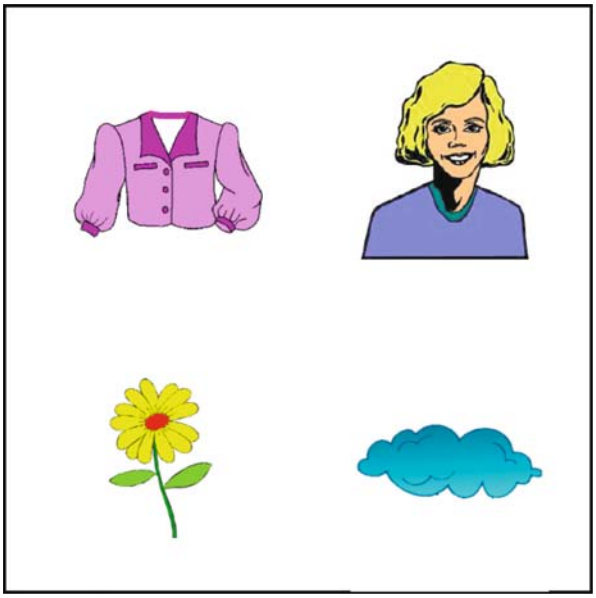

Fig. 4 Visual context for spoken sentences: *Die Frau bügelt/sieht die Bluse*

the high-frequency phonological distractor *Blume* should draw more looks than the low frequency target while hearing the ambiguous part of the target (e.g., /blu/ in "Bluse").

Not surprisingly, Weber and Crocker [36] replicated the finding that when the verb was not semantically constraining the set of potential object arguments, German listeners fixated both the picture of the target and the picture of the phonological distractor during the ambiguous part of the target. Thus, both *Bluse* and *Blume* were considered as potential object arguments following the unrestrictive verb *siehst*. No activation of the phonological distractor was, however, observed when the preceding verb was excluding the distractor as a likely object referent (see Fig. 5); looks went almost exclusively to the target *Bluse* following the restrictive verb *bügelt*. At first glance, it clearly seems that semantic information provided by the verb is sufficient to exclude semantically inappropriate distractors even when they are high in lexical frequency. However, this complete lack of distractor activation in semantically constraining context should be taken with some caution; lexical frequency was predicted after all to make the phonological distractor more attractive than the target, albeit only when there is no semantic restriction on the target. This was, however, not what Weber and Crocker [36] found; rather the picture of the target and the phonological distractor were equally attractive in unrestrictive sentences. This seems surprising given the earlier findings of lexical frequency effects in eye tracking (see [10]).

In contrast to these earlier studies, the German participants in [36] had no specific task during the experiment, other than to listen to the speech and to look at

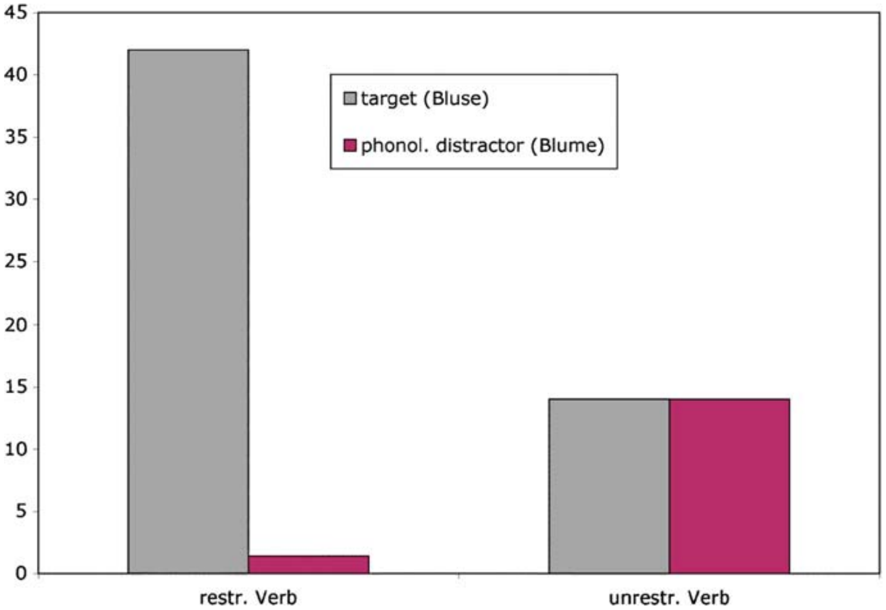

Fig. 5 Attractiveness of target and phonological distractor between 300 and 600 ms after target onset, measured as added percentages of looks over unrelated distractor. No specific task

the screen. Previously, targets in lexical frequency studies had been presented in semantically empty carrier phrases which simply instructed participants to click on a displayed object (e.g., click on the blouse). In a second experiment, Weber and Crocker [36] therefore tested whether, in combination with a task, lexical frequency effects could be observed with their materials. They presented the same materials to a new set of German listeners, the only difference being that participants were told to click on the picture of the second argument in the sentence. This time, the phonological distractor *Blume* was indeed more attractive than the target *Bluse* in unrestrictive sentences (see Fig. 6). Just by having an explicit task, lexical frequency effects emerged. This dominance of high-frequency phonological distractors is therefore consistent with previous studies on lexical frequency effects that employ a click task.

But even more interesting for the question of semantic information, Weber and Crocker [36] found activation of the phonological distractor in semantically constraining contexts; even though the verb information in *bügelt* should have rendered the phonological distractor *Blume* an unlikely candidate for the object argument, German listeners still look at it more than would be expected. In the constraining sentences, the target was overall more attractive than the phonological distractor, but the phonological distractors drew also a considerable proportion of looks. This suggests that effects of preceding verb information can indeed be modulated by lexical frequency.

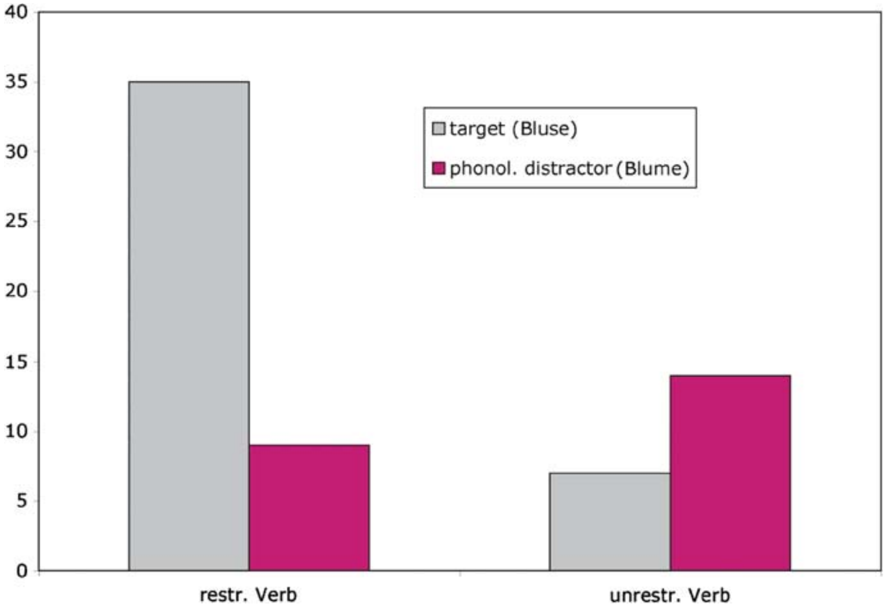

Fig. 6 Attractiveness of target and phonological distractor between 300 and 600 ms after target onset, measured as added percentages of looks over unrelated distractor. Clicking task

The fact that activation of semantically inappropriate, high-frequency distractors was only found when the participants' task was to click on the last argument in the sentence, suggests that frequency effects in eye tracking are sensitive to task specific demands. Apparently, only when listeners' attention is purposefully directed to the verb arguments, are frequency effects observable. This finding speaks for a human parser that is not only applying different information sources incrementally, but that is also sensitive to cognitive task demands.

2.4 Semantic Information and Visual Context

We have observed above that situated language comprehension rapidly directs visual attention in a relevant scene, both to mentioned and anticipated referents. An important question about these findings is the extent to which they are indicative of general comprehension mechanisms, or whether the scene objects themselves contribute to the forming of specific expectations for verb arguments. Weber and Crocker [37] therefore investigated further the influence of visual context on constraining verb information in a cross-modal priming experiment.

Lexical decision times are known to be faster following semantically related objects than semantically unrelated objects; that is, listeners respond faster to *nurse* after *doctor* than after *grass* (e.g., [28]). Also verbs have been shown to prime typical agents, patients, and instruments (e.g., [12]). In a first step, Weber and Crocker

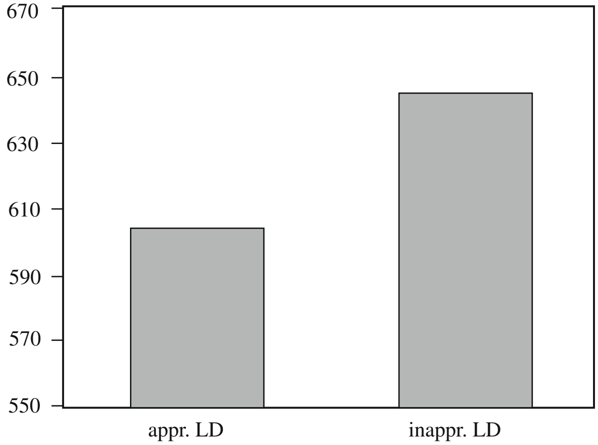

Fig. 7 Average lexical decision times for semantically appropriate and inappropriate object arguments. Only auditory prime

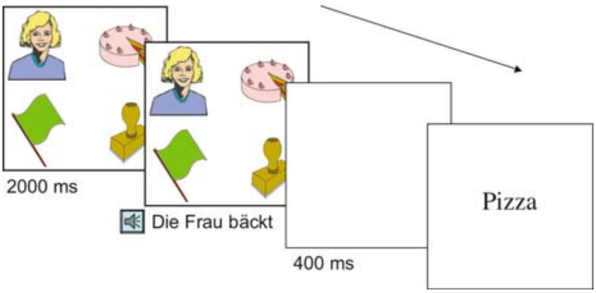

Fig. 8 Example for a trial with a combination of auditory and visual primes

[37] replicated this finding for German with a simple cross-modal priming study in which selectional verb information could prime object arguments. In their study, German listeners were presented with sentence onsets which had a semantically restrictive verb (e.g., *Die Frau bäckt*, “the woman bakes”); a lexical decision task for visually presented nouns followed the auditory prime sentence fragment. The visual lexical decision items were either semantically appropriate as arguments for the verb (e.g., *Pizza*, “pizza”) or inappropriate (e.g., *Palme*, “palm tree”) as arguments for the verb. As expected, reaction times were faster for semantically appropriate items than for inappropriate ones (see Fig. 7), replicating the well-known semantic priming effect for German.

In order to further investigate the influence of the visual context on forming expectations about upcoming verb arguments, Weber and Crocker displayed in a second study objects on a screen, simultaneously with the auditory prime (see Fig. 8). The displays were typical for eye-tracking studies and showed four objects: the agent of the sentence onset (e.g., *die Frau*, “the woman”), an object either

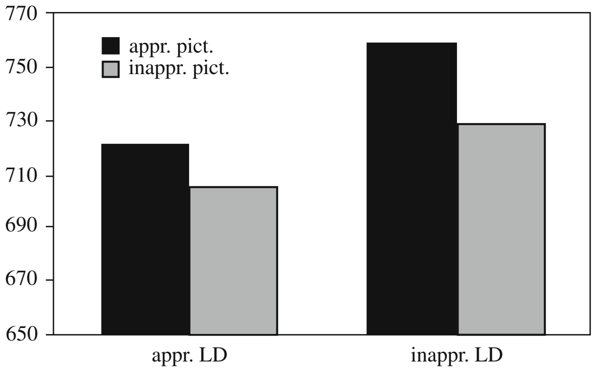

Fig. 9 Average lexical decision times for semantically appropriate and inappropriate object arguments. Auditory and visual prime

semantically appropriate as argument for the verb (e.g., *Torte*, "pie") or inappropriate (e.g., *Tanne*, "pine") and two distractor objects.

As before, a lexical decision task to visually presented nouns followed the primes. Both the appropriate visual argument (e.g., *Torte*, "pie") and the appropriate lexical decision item (e.g., *Pizza*, "pizza") were highly plausible arguments for the sentence onsets (as defined by a rating study). As in the first study, reaction times were faster for lexical decision items which were semantically appropriate than for items which were inappropriate (see Fig. 9). Surprisingly, however, reaction times were slowed when the display included a picture of an appropriate argument prior to lexical decision. This semantic interference (rather than facilitation) from appropriate pictures occurred both when lexical decision items were appropriate and when they were inappropriate. Thus, visual context did influence reaction times in the sense that it gave competition to the lexical decision items. On the other hand, facilitated lexical decision times for appropriate items, regardless of the scene, provide evidence for a purely linguistic anticipation of upcoming verb arguments (confirming the gated completion findings of [2]). We suggest that visually attending the picture of an appropriate object based on supporting auditory input leads to contextually grounded expectations concerning which object would follow as the verb argument; when the visually supported expectations were not met by the target word, lexical decision times were slowed across the board.

2.5 The Influence of the Scene: Depicted Events and Their Priority

The studies described above provide diverse evidence supporting the notion that the human language comprehension system is able to rapidly adapt to, and exploit, a range of linguistic information sources: discourse context, prosody, lexical frequency, and verb semantics. We further noted that anticipatory inspection of relevant depicted referents not only reflects incremental interpretation and expectations, but further instantiates those expectations with the depicted object. A natural question in the case of situated language processing, therefore, is whether more complex scene information can influence spoken language understanding. Previous work by Tanenhaus and colleagues [35] has shown, for example, the rapid influence of visual referential context on ambiguity resolution in on-line situated utterance processing. Listeners were presented with a scene showing either a single apple or two apples, and the utterance *Put the apple on the towel in the box*. Eye-movements revealed that the interpretation of the phrase *on the towel*, as either the location of the apple versus its desired destination was influenced by the visual context manipulation. Sedivy et al. [33] further demonstrated the influence of a visual referential contrast: listeners looked at a target referent (e.g. the tall glass) more quickly when the visual context displayed a contrasting object of the same category (a small glass) than when it did not.

An eye-tracking study by Knoeferle et al. [21] investigated the interpretation of German SVO and OVS sentences with case-ambiguous initial NPs. Structural disambiguation took place only at a second NP that was clearly case marked as either nominative or accusative (e.g., *die Prinzessin malt offensichtlich den Fechter/der Pirat*, "the princess$_{NOM,ACC}$ paints apparently the fencer$_{ACC}$/the pirat$_{NOM}$"). In the accompanying scenes, however, depicted actions were potentially able to resolve the ambiguity as soon as the verb was encountered (see Fig. 10). Their findings revealed anticipatory post-verbal eye movements to the appropriate second argument based on verb-mediated identification of the relevant scene event, and crucially before the disambiguating second NP was heard. The time-course and pattern of gaze clearly suggest that listeners were able to use depicted events to resolve the ambiguity and assign grammatical functions appropriately, just as they have been shown to use linguistic [18] and prosodic [38] constraints.

Given that information sources as diverse as syntax, semantic, intonation, and depicted events can so rapidly and effectively be used during situated spoken

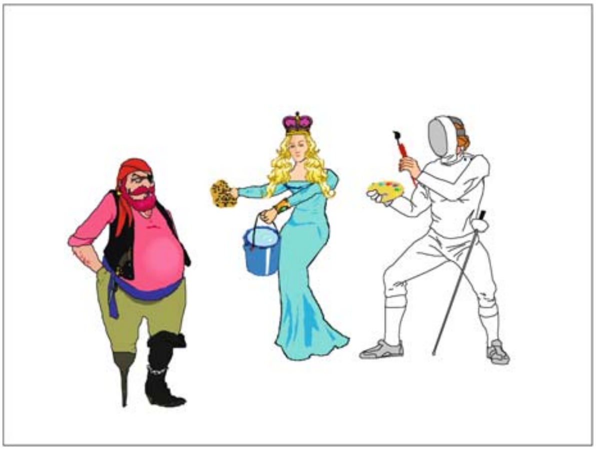

Fig. 10 Visual context for spoken sentences: *Die Princessin wäscht/malt gerade . . .*

language comprehension, Knoeferle and Crocker [20] investigated the time course with which world knowledge about typical events [2] and information from the atypical scene events interacted [21], and, crucially, the *relative importance* of these information sources. In a German eye-tracking study, they investigated the anticipation of both stereotypical role-fillers, based on verb expectations, and depicted role-fillers, based on depicted events in syntactically unambiguous sentences.

Their findings confirmed, in a single study, that people are able to rapidly and equally exploit both information sources, linguistic or visual, when either kind of constraining information is available to anticipate thematic role fillers. Crucially, however, when they pitted the two information sources against each other, they observed a greater relative importance of verb-mediated depicted events information over stereotypical thematic role knowledge associate with the verb. When listeners heard a sentence beginning *Den Pilot bespitzelt gleich* . . . , the verb (e.g., *bespitzelt*) (spies-on) identifies two different agents on a scene as relevant, participants prefer upcoming agents that match with a displayed action (e.g., *a wizard*) over agents that match with their world knowledge (e.g., *a spy*) (see Fig. 11). Eye movements to the agent depicting the action of the verb occur shortly after the verb and crucially before the agent was mentioned in the utterance.

To further investigate the priority and use of scene events, Knoeferle and Crocker [22] conducted a series of experiments investigating the temporal interdependency between *dynamic* visual context and utterance comprehension. Exploiting the "blank screen paradigm", event scenes were presented prior to the onset of an utterance and then replaced by a blank screen either before or during the utterance. Additionally, two of the experiments featured scenes involving dynamic events, i.e., actions were depicted as occurring over time, introducing an aspectual dimension to the depicted events, which were furthermore coupled with verb and adverb tense manipulations in the utterances used in the third experiment. The findings suggested that people do use scene event information even when it is no longer

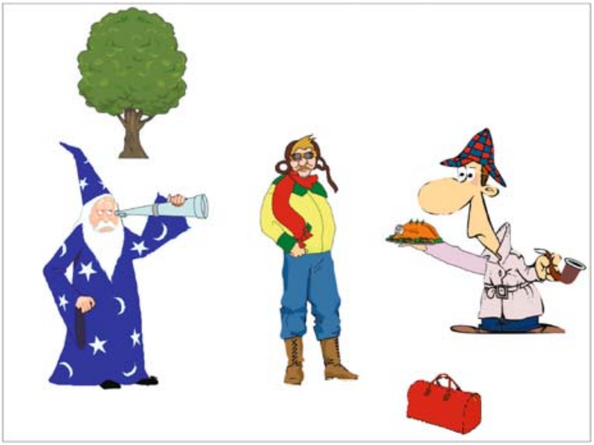

Fig. 11 Visual context for spoken sentences: *Den Pilot bespitzelt*. . .

present, but that the relative priority with respect to other information sources is strongest when events are co-present, and may decay over time.

To account for both the rapid interaction of linguistic and visual information and the observed preference for the information from depicted events, Knoeferle and Crocker [20] posit the Coordinated Interplay Account (CIA), which outlines how the unfolding utterance guides attention in the visual scene to establish reference to objects and events. Once these are identified, the attended information rapidly influences comprehension of the utterance, allowing the anticipation of upcoming arguments not yet mentioned by virtue of their relationship to the objects and events thus established. The close temporal interaction between comprehension and attention in the scene is suggested as the principal reason for the relative priority of the immediately depicted information over stereotypical knowledge in situated comprehension. Knoeferle and Crocker [20] conjecture that there may be a development basis for this preference, arising from the important role that the immediate environment plays as a child learns to ground concepts with visual referents during language acquisition. Mayberry et al. [26] have furthermore developed a connectionist model which instantiates the CIA. The architecture, described in detail in the Chapter by Mayberry and Crocker (this volume), models many of the findings described above, including the priority of scene event information.

In more recent work, Knoeferle and Crocker [22] further refine the CIA to incorporate a working memory (WM) component that contains the current interpretation of an utterance, expectations based on linguistic and world knowledge, and information from objects and events in a dynamic scene (Fig. 12). In order to explain the reduced priority of events that are no longer co-present, the account postulates that

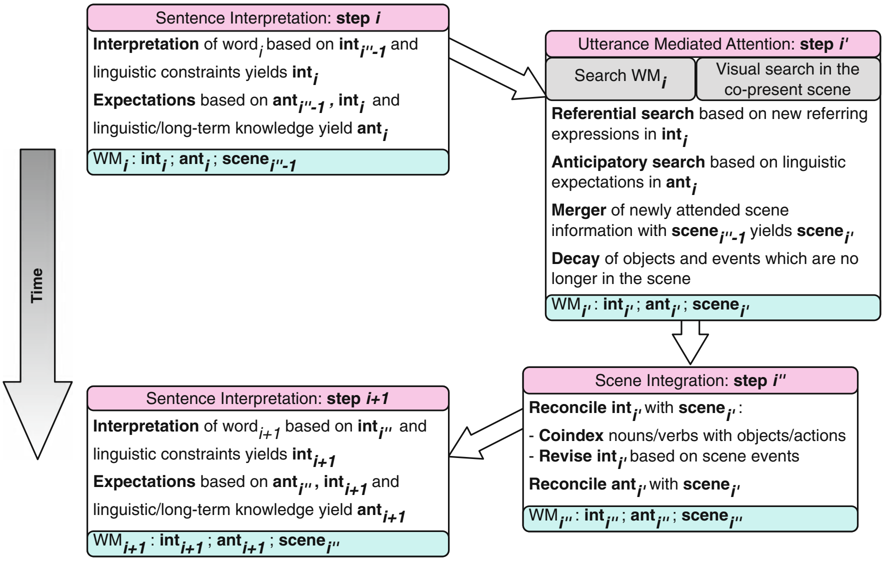
Sentence Interpretation: step i
Interpretation of word$_i$ based on int$_{i''-1}$ and linguistic constraints yields int$_i$
Expectations based on ant$_{i''-1}$, int$_i$ and linguistic/long-term knowledge yield ant$_i$
WM$_i$: int$_i$; ant$_i$; scene$_{i''-1}$
Utterance Mediated Attention: step i'
Search WM$_i$
Visual search in the co-present scene
Referential search based on new referring expressions in int$_i$
Anticipatory search based on linguistic expectations in ant$_i$
Merger of newly attended scene information with scene$_{i''-1}$ yields scene$_{i'}$
Decay of objects and events which are no longer in the scene
WM$_{i'}$: int$_{i'}$; ant$_{i'}$; scene$_{i'}$
Time
Scene Integration: step i''
Reconcile int$_{i'}$ with scene$_{i'}$:
- Coindex nouns/verbs with objects/actions
- Revise int$_{i'}$ based on scene events
Reconcile ant$_{i'}$ with scene$_{i'}$
WM$_{i''}$: int$_{i''}$; ant$_{i''}$; scene$_{i''}$
Sentence Interpretation: step $i+1$
Interpretation of word$_{i+1}$ based on int$_{i''}$ and linguistic constraints yields int$_{i+1}$
Expectations based on ant$_{i''}$, int$_{i+1}$ and linguistic/long-term knowledge yield ant$_{i+1}$
WM$_{i+1}$: int$_{i+1}$; ant$_{i+1}$; scene$_{i''}$

Fig. 12 The Coordinated Interplay Account (CIA): The time course of processes, informational dependencies, and working memory in situated spoken language comprehension [22]

items in working memory decay with time, affecting their influence on the developing interpretation of the unfolding utterance. The introduction of working memory into the CIA in which the accessibility of representations of scene objects and events are dependent on their decaying activation provides a reasonable explanation for the observed effects that is also in accord with current theories of sentence comprehension (see Lewis and Vasishth [23] for discussion as well as [3] for a broader view of the role of working memory in cognition).

3 Conclusions

Human sentence processing is not only incremental but also anticipatory: upon encountering the initial words of a sentence, not only do people immediately begin constructing detailed interpretations, they also initiate hypotheses or expectations about what is likely to follow. The empirical research of the ALPHA project focussed mainly on two aspects of such anticipatory behavior. In the first phase of the project the emphasis of empirical research lay in establishing the information sources which contribute to incremental interpretation and anticipation of forthcoming arguments. It was found that such sources include morphosyntactic and lexical verb information (e.g., [18]), world-knowledge (e.g., [29]), and information from the visual context (e.g., [21]). In the second phase of the project, the focus of empirical research was to determine the extent to which initial interpretation preferences are influenced by different short-term and long-term constraints such as lexical frequency, linguistic context, visual context, and prosodic information. This chapter has highlighted some of the most important empirical findings from the second phase.

While long-term exposure to language can result, for instance, in preferences for certain syntactic structures, biases for lexical verb frames, and frequency effects for lexical choices, recent linguistic and visual context is also exploited on-line to influence understanding. Given the diverse nature of long-term and short-term constraints it seems possible that they affect the comprehension processes differently: for instance, one type of constraint could be dominant. It would, for instance, appear plausible that long-term knowledge derived from experience with language and the world is always accorded greater weight than short-term constraints. Long-term constraints are presumably *routinized* within the processing mechanisms, while short-term constraints may be more variable depending on the specifics of the communicative situation and task. Alternatively, the *here and now* relevance of a communicative situation may foreground short-term constraints in the immediate (linguistic and non-linguistic) context over what we know based on our long-term experience. The first account is appealing since rapid and preferred reliance on long-term experience would enable efficient processing because such long-term knowledge is readily available from memory. On the other hand, an ever-changing dynamic environment and the necessity of adapting to different communicative situations and tasks would appear to favor the second account, placing emphasis on the use of short-term contextual constraints.

Consider our findings in the light of these two accounts. On the one hand, they confirm that long-term biases (e.g., structural bias towards SVO) cannot be fully overridden by short-term contextual constraints that are linguistic in nature: While both information about grammatical function in preceding context (see Sect. 2.1) and prosodic marking of object-first sentence (see Sect. 2.2) could weaken the processing difficulties usually encountered with object-first structure, these short-term constraints were not sufficient to fully generate the interpretation of an object-first sentence. Similarly, lexical frequency could modulate, but definitely not fully change the expectations for an object argument based on restrictive verb information (see Sect. 2.3), and also the results from Sect. 2.4 speak clearly for an interplay of both the visual context and the verb information.

On the other hand, short-term constraints arising from depicted event information appear to dominate (and not just modulate) the stereotypical knowledge of the actions an agent performs (see [20]). This finding together with the strong influence of depicted events on structural disambiguation of locally structurally ambiguous German utterances (see [21]) suggests an account of situated language comprehension in the tradition of the Coordinated Interplay Account posited by Knoeferle and Crocker [20, 22]. In situated comprehension situations, information from the immediate visual context, at least when it depicts role relations between event participants, is accorded great importance for on-line language comprehension.

An added dimension with respect to the use of short- and long-term constraints comes from the presented effects of task: lexical frequency only biased the anticipation of a visually presented object when the task was to click on the target object (Sect. 2.3). The studies in Sect. 2.4 further revealed an interesting combination of long- and short-term constraints: While we observed clear support for the general anticipation of objects based on verb-derived expectations, the scene then instantiated these expectations causing interference with the lexical decision targets that did not match objects in the scene when these general expectations identified a plausible referent in the scene. The pattern of observations is consistent with the Coordinated Interplay Account, which generally argues for the influence of scene information once it has been identified by the utterance as relevant, typically through explicit or anticipated reference to scene objects or events. Taken together, the empirical findings of the ALPHA project speak for a human sentence comprehension system that rapidly integrates diverse informational constraints, derived from both long-term experience and the immediate context, and weighs them depending on the situation and task.

References

1. Altmann, G., Garnham, A., Dennis, Y. Avoiding the garden path: Eye movements in context. Journal of Memory and Language, 31:685–712 (1992).
2. Altmann, G.T.M., Kamide, Y. Incremental interpretation at verbs: Restricting the domain of subsequent reference. Cognition, 73:247–264 (1999).
3. Baddeley, A.D. Working Memory. Oxford, UK; New York: Oxford University Press (1986).

4. Branigan, H., Pickering, M., Liversedge, S. Syntactic priming: investigating the mental representation of language. Journal of Psycholinguistic Research, 24:489–506 (1995).
5. Chater, N., Crocker, M.W., Pickering, M. The rational analysis of inquiry: The case for parsing. In N. Chater, M. Oaksford (Eds.), Rational Analysis of Cognition (pp. 441–468). Oxford, UK; New York: Oxford University Press (1998).
6. Crocker, M., Keller, F. Probabilistic grammars as models of gradience in language processing. In G. Fanselow et al. (Ed.), Gradience in Grammar: Generative Perspectives (pp. 227–245). Oxford, UK; New York: Oxford University Press (2006).
7. Crocker, M.W. Computational Psycholinguistics: An Interdisciplinary Approach to the Study of Language. Dordrecht: Kluwer (1996).
8. Crocker, M.W. Rational models of comprehension: Addressing the performance paradox. In A. Cutler (Ed.), Twenty-First Century Psycholinguistics: Four Cornerstones (pp. 363–380). Hillsdale, NJ: Lawrence Erlbaum Associates (2005).
9. Crocker, M.W., Brants, T. Wide-coverage probabilistic sentence processing. Journal of Psycholinguistic Research, 29(6):647–669 (2000).
10. Dahan, D., Magnuson, J., Tanenhaus, M. Time course of frequency effects in spoken-word recognition: Evidence from eye movements. Cognitive Psychology, 42:361–367 (2001).
11. Elman, J.L. Finding structure in time. Cognition Science, 14(2):179–211 (1990).
12. Ferretti, T., McRae, K., Hatherell, A. Integrating verbs, situation schemas, and thematic role concepts. Journal of Memory and Language, 44:516–547 (2001).
13. Frazier, L., Fodor, J. The sausage machine: A new two-stage parsing model. Cognition, 6:291–325 (1978).
14. Garnsey, S., Pearlmutter, N., Myers, E., Lotocky, M. The contribution of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. Journal of Memory and Language, 37:58–93 (1997).
15. Gibson, E. Linguistic complexity: Locality of syntactic dependencies. Cognition, 68:1–76 (1998).
16. Grice, M., Baumann, S. Deutsche intonation und GToBI. Linguistische Berichte, 191: 267–298 (2002).
17. Hemforth, B. Kognitives Parsing: Repräsentation und Verarbeitung Sprachlichen Wissens. Sankt Augustin: Infix-Verlag (1993).
18. Kamide, Y., Scheepers, C., Altmann, G. Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. Journal of Psycholinguistic Research, 32:37–55 (2003).
19. Kjeelgaard, M., Speer, S. Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. Journal of Memory and Language, 40:153–194 (1999).
20. Knoeferle, P., Crocker, M. The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. Cognitive Science, 30:481–529 (2006).
21. Knoeferle, P., Crocker, M., Scheepers, C., Pickering, M. The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye movements in depicted events. Cognition, 95:95–127 (2005).
22. Knoeferle, P., Crocker, M.W. The influence of recent scene events on spoken comprehension: Evidence from eye movements. Journal of Memory and Language (Special Issue on Language-Vision Interaction), 57(2):519–543 (2007).
23. Lewis, R.L., Vasishth, S., Dyke, J.A.V. Computational principles of working memory in sentence comprehension. Trends in Cognitive Science, 10:447–454 (2006).
24. Marslen-Wilson, W. Linguistic structure and speech shadowing at very short latencies. Nature, 244:522–523 (1973).
25. Marslen-Wilson, W. Activation, competition, and frequency in lexical access. In G. Altmann (ed.), Cognitive Models of Speech Processing (pp. 148–172). Cambridge, MA: MIT Press (1990).
26. Mayberry, M., Crocker, M., Knoeferle, P.A connectionist model of the coordinated interplay of scene, utterance, and world knowledge. In 28th Annual Conference of the Cognitive Science Society. Vancouver, Canada (2006).

27. McRae, K., Spivey-Knowlton, M., Tanenhaus, M. Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. Journal of Memory and Language, 38:283–312 (1998).
28. Meyer, D., Schvaneveldt, R. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. Journal of Experimental Psychology, 90:227–234 (1971).
29. Muckel, S., Scheepers, C., Crocker, M., Muller, K. Anticipating German particle verb meanings: Effects of lexical frequency and plausibility. In 8th Annual Conference on Architectures and Mechanisms for Language Processing. Tenerife, Spain (2002).
30. Pickering, M., Traxler, M., Crocker, M. Ambiguity resolution in sentence processing: Evidence against likelihood. Journal of Memory and Language, 43:447–475 (2000).
31. Rayner, K., Carlson, M., Frazier, L. The interaction of syntax and semantics during sentence processing. Journal of Verbal Learning and Verbal Behavior, 22:358–374 (1983).
32. Rayner, K., Raney, G.E. Eye movement control in reading and visual search effects of word frequency. Psychonomic Bulletin & Review, 3:245–248 (1996).
33. Sedivy, J.C., Tanenhaus, M.K., Chambers, C.G., Carlson, G.N. Achieving incremental semantic interpretation through contextual representation. Cognition, 71:109–148 (1999).
34. Swets, B., Desmet, T., Clifton, C., Ferreira, F. Underspecification of syntactic ambiguities: Evidence from self-paced reading. Memory & Cognition, 36:201–216 (2008).
35. Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., Sedivy, J.C. Integration of visual and linguistic information in spoken language comprehension. Science, 268:1632–1634 (1995).
36. Weber, A., Crocker, M. Top-down anticipation versus bottom-up lexical access: Which dominates eye movements in visual scenes? In 19th Annual CUNY Conference On Human Sentence Processing. New York City, New York (2006).
37. Weber, A., Crocker, M. The influence of the scene on linguistic expectations: Evidence from cross-modal priming in visual worlds. In 20th Annual CUNY Conference On Human Sentence Processing. New York City, New York (2007).
38. Weber, A., Grice, M., Crocker, M. The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. Cognition, 99:B63–B72 (2006).
39. Weber, A., Müller, K. Word order variation in German main clauses: A corpus analysis. In Proceedings of the 20th International Conference on Computational Linguistics (pp. 71–77). Geneva (2004).
40. Weber, A., Neu, J. Assignment of grammatical functions in discourse context and word-order ambiguity resolution. In 16th Annual CUNY Conference On Human Sentence Processing. Boston, Massachusetts (2003).