Expectation and Locality Effects in German Verb-final Structures

Roger P. Levy[a], Frank Keller[b]

[a] *Department of Linguistics, UC San Diego, 9500 Gilman Drive #0108, La Jolla, CA 92093-0108, USA, phone: +1 (858) 534-7219, fax: +1 (858) 534-4789*

[b] *School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK, phone: +44 131 650 4407, fax: +44 131 650 6626*

Abstract

Probabilistic expectations and memory limitations are central factors governing the real-time comprehension of natural language, but how the two factors interact remains poorly understood. One respect in which the two factors have come into theoretical conflict is the documentation of both *locality* effects, in which more dependents preceding a governing verb increase processing difficulty at the verb, and *anti-locality* effects, in which more preceding dependents facilitate processing at the verb. However, no controlled study has previously demonstrated both locality and anti-locality effects in the same type of dependency relation within the same language. Additionally, many previous demonstrations of anti-locality effects have been potentially confounded with lexical identity, plausibility, and sentence position. Here, we provide new evidence of both locality and anti-locality effects in the same type of dependency relation in a single language—verb-final constructions in German—while controlling for lexical identity, plausibility, and sentence position. In main clauses, we find clear anti-locality effects, with the presence of a preceding dative argument facilitating processing at the final verb; in subject-extracted relative clauses with identical linear ordering of verbal dependents, we find both anti-locality and locality effects, with processing facilitated when the verb is preceded by a dative argument alone, but hindered when the verb is preceded by both the dative argument and an adjunct. These results indicate that both expectations and memory limitations need to be accounted for in any complete theory of online syntactic comprehension.

Introduction

A large body of experimental evidence in psycholinguistics indicates that the human sentence processor is able to build up expectations about upcoming linguistic material based on the input it has received so far, and that these expectations can influence both real-time comprehension behavior and its neural correlates (Ehrlich & Rayner, 1981; Kutas & Hillyard, 1980, 1984; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995; Altmann & Kamide, 1999). Although earlier work documenting these effects focused primarily on expectations at the lexical level, more recent work has provided evidence for expectations at the level of syntactic constituency on the basis of grammatical analysis of prior linguistic content (Lau, Stroud, Plesch & Phillips, 2006; Staub & Clifton, 2006; Staub, Clifton & Frazier, 2006; Jaeger, Fedorenko, Hofmeister & Gibson, 2008; Levy, Fedorenko, Breen & Gibson, 2012).

A particularly clear example of how online processing can be sharply modulated by fine-grained differences in the grammatical structure of preceding context is provided by Konieczny & Döring (2003), who investigated verb-final structures in German such as (1) below:

(1) a. Die Einsicht, dass der Freund *des* Kunden das Auto aus *Plastik/Freude* verkaufte,
The insight, that the friend *the.*GEN customer the automobile out_of *plastic/joy* sold,
erheiterte die Anderen.
amused the others.
"The insight that the friend of the customer sold the automobile {made of plastic/out of joy} amused the others."

b. Die Einsicht, dass der Freund *dem* Kunden das Auto aus *Plastik/Freude* verkaufte,
The insight, that the friend *the.*DAT customer the automobile out_of *plastic/joy* sold,

erheiterte die Anderen.
amused the others.
"The insight that the friend sold the customer the automobile {made of plastic/out of joy} amused the others."

The preceding contexts in (1-a) and (1-b) differ by only a change in a single character—*s* versus *m*—but this difference dramatically changes the grammatical structure of the sentence: in (1-a), *des Kunden* is a genitive postmodifier of the noun *Freund*, whereas in (1-b), *dem Kunden* is a dative dependent of the subordinate verb *verkaufte*. Intuitively, encountering a preverbal dative constrains the argument structure of the yet-to-be-seen subordinate verb, which sharpens expectations about both *when* the verb will appear and *what* it will turn out to be when it appears (Konieczny, 1996; Levy, 2008a). Konieczny & Döring (2003) found evidence supportive of this intuition in a free-reading eye-tracking experiment: regression-path durations (the time elapsed between first fixation on a word and the first fixation beyond it) were shorter for *verkaufte* in the verbal-dependent condition (1-b) than in the nominal-dependent condition (1-a). However, it was not simply the *presence* of more preverbal dependents that triggered this expectation-based facilitation: Konieczny & Döring also manipulated the *type* of PP immediately preceding the subordinate verb, comparing a PP post-modifying the preceding NP, such as *aus Plastik*, with a preverbal PP modifier, such as *aus Freude*. This manipulation did not have a significant effect on reading behavior. Konieczny & Döring's 2003 study therefore provides evidence for expectation-based facilitation, but only derived from complements of the verb, not for modifiers such as PP adjuncts which intuitively place much less constraint on the governing verb than complements do.

Such effects are sometimes referred to as *anti-locality* effects, because they run contrary to common assumptions regarding the processing cost incurred when the sentence processor has to complete dependency relations between previous input and the word currently processed—in the case of (1), the dependencies between the verb *verkaufte* and its subject *der Freund* and object *das Auto*, plus the indirect object *dem Kunden* and/or the adjunct *aus Freude* in the appropriate conditions. The theories most closely associated with these notions of dependency-completion processing cost are Dependency Locality Theory (DLT, Gibson 2000; Grodner & Gibson 2005; also known as Syntactic Prediction Locality Theory, Gibson 1998) and Similarity-based Interference (SBI, Gordon, Hendrick & Johnson 2001, 2004; Gordon, Hendrick, Johnson & Lee 2006; Lewis & Vasishth 2005; Lewis, Vasishth & Van Dyke 2006; Van Dyke & Lewis 2003). In DLT, dependency completion involves an *integration cost* determined by the number and distance of the dependents that precede the current word. In SBI, retrieval of preceding dependents is a precondition for dependency completion, and this retrieval process is subject to *interference* from similar constituents elsewhere in previous input. To go into slightly greater detail regarding these dependency completion cost theories: in standard DLT theory, the integration cost for each preceding dependent is equal to the number of discourse referents (effectively, the number of NPs) intervening between the dependent and the current word; and total integration cost is the sum of integration costs across all previous dependents.[1] In (1-b), for example, the integration cost at *verkaufte* would be 2 units greater than in (1-a), since *dem Kunden* is a preverbal dependent only in (1-b), and two discourse referents intervene between it and the clause-final verb. Both theories thus predict locality effects for examples like (1): adding the dative NP should make processing more difficult, rather than easier, given that an additional dependent has to be integrated at the subordinate verb.

Substantial evidence has been adduced for locality effects in English (see Gibson 1998, for an overview), and more recent evidence has suggested the presence of strong locality effects in the processing of Chinese relative clauses (Hsiao & Gibson, 2003) and Russian relative clauses (Levy, Fedorenko & Gibson, 2011). However, locality effects have been elusive in many other languages, including German, Japanese, and Hindi, where *anti-locality* effects of the type found by Konieczny & Döring have been reported (Konieczny, 2000; Nakatani & Gibson, 2008; Vasishth & Lewis, 2006): adding a preverbal dependent facilitates rather than hinders processing at the clause-final verb. A locality-based interpretation of the Konieczny & Döring results might be that the traditional method of quantifying integration cost is wrong. For example, perhaps the total integration cost at a final verb should be taken to be the *maximum* of the integration costs of each preceding dependent, rather than the *sum*—a measure that would be natural

[1] The predictions of SBI would be qualitatively similar in the cases examined in this paper, but precise quantification is more complex due to its dependence on the similarity-space representations of all constituents involved. For simplicity, we therefore use DLT integration costs to exemplify predictions that we would expect to hold of both theories.

if integration of preceding dependents were assumed to occur in parallel rather than serially. On this view, the total integration costs in (1-a) and (1-b) should be identical, since *der Freund* incurs the highest integration cost of any preverbal dependent at its governing verb, and is separated from it by the same number of discourse referents. However, the data of Konieczny (2000) speak against this possibility. Among other conditions, Konieczny had participants read sentences of the following form:

(2) a. Er hat die Rose hingelegt, und ...
He has the rose laid_down, and ...
"He put down the rose, and ..."

b. Er hat die Rose *auf den Tisch* gelegt, und ...
He has the rose on the table laid, and ...
"He put the rose on the table, and ..."

Here, *die Rose* (or *Er*, if it is considered dependent on the participle *(hin)gelegt* rather than on the auxiliary *hat*) is the most distant dependent from the final participial verb *(hin)gelegt*, and is separated from the governing verb by more discourse referents in (2-b) than in (2-a). Hence both maximum and total integration costs are higher in (2-b) than in (2-a). Contra the predictions of DLT and SBI, Konieczny found shorter reading times on the final participial verb in (2-b) than in (2-a). However, there are several crucial confounds in Konieczny's 2000 study: the participial verb's position within the sentence and the identity of the immediately preceding word—both of which are believed to affect reading times (Mitchell, 1984; Ferreira & Henderson, 1993)—vary across conditions, and in many items, such as the one in (2), the participial verb itself varied across condition as well. Furthermore, neither the Konieczny (2000) nor the Konieczny & Döring (2003) studies controlled for sentence plausibility, which also are known to affect reading times under some circumstances (Traxler & Pickering, 1996; Ni, Crain & Shankweiler, 1996; Garnsey, Perlmutter, Meyers & Lotocky, 1997; Rayner, Warren, Juhasz & Liversedge, 2004).

A conclusive account of the presence or absence of locality effects in German is thus still lacking, but would be of considerable interest given the combination of locality and anti-locality results across multiple experiments in a variety of languages. There have been some recent signs that even within a single language, an adequate sentence-processing mechanism must be able to account for both locality and anti-locality effects in verbal processing: Jaeger et al. (2008) demonstrate anti-locality effects at verbs in English, and Demberg & Keller (2008a) found evidence for both locality and expectation effects in an eye-tracking corpus of English newspaper text. However, it has not been previously demonstrated in controlled experiments that both locality and anti-locality effects can arise in a single syntactic dependency configuration in a single language. Such a finding would provide even stronger evidence that a complete model of human sentence processing must be able to account for both types of effects. We report such a finding in this paper, showing that both locality and expectation effects can be found in clause-final verbs in German. The key difference between our results and previous work on German verbal processing is that we find that locality effects are only detectable at much higher levels of memory load than have previously been studied. In this situation, both locality and anti-locality effects may manifest themselves simultaneously.[2]

Experiment 1

The design of our first experiment is qualitatively similar to that of Konieczny (2000) and Konieczny & Döring (2003)—investigating the effects of adding preverbal dependents varying in predictive value to a verb-final clause—but eliminates several confounds present in previous work, including identity of the critical verb and the words immediately preceding it, position of the critical verb within the sentence, and plausibility. Furthermore, we constructed our materials so that memory loads would be much higher than has been used in previous experiments on verb-final processing in any language, intuitively maximizing the opportunity of finding locality effects. The materials of Konieczny & Döring (2003), for example (see (1)), did not include a strong memory load manipulation: by the DLT metric (Gibson, 2000), the dative NP *dem Kunden* incurs an integration cost of two at the main verb, as there are two discourse

[2] Since first submission of this work we have become aware of recent work by Vasishth & Drenhaus (2011), who also present evidence suggesting locality effects on German verbs. These empirical results complement our own, and our theoretical interpretation is broadly compatible with that of Vasishth & Drenhaus.

referents intervening between the dative NP and the verb (*das Auto* and *aus Freude* or *aus Plastik*). Furthermore, the integration of the verb-modifying adjunct *aus Freude* incurs an integration cost of zero, as it is directly adjacent to the verb it needs to be integrated with, resulting in a distance of zero.[3] It follows that the dative and the adjunct NP in (1) only trigger a relatively small change in integration cost (two additional units); this could be the reason for the absence of locality effects in Konieczny & Döring's 2003 study. It seems possible that locality effects might become visible in an experiment that increased the number of intervening discourse referents more drastically.[4] Our design picks up on this idea, and can be illustrated using the simplified examples in (3). We manipulated two factors: the presence of a dative NP such as *dem Sohn* ("the.DAT son") in the subordinate clause, and the presence of a PP adjunct modifying the verb such as *zur Ahndung* ("as payback"). The resulting sentence contains neither the dative nor the adjunct as in (3-a), or just the adjunct or just the dative as in (3-b) and (3-c), or both as in (3-d).

(3) a. Hans hat den Fußball versteckt.
Hans has the.ACC football hidden.
"Hans hid the football."

b. Hans hat zur Ahndung den Fußball versteckt.
Hans has as payback the.ACC football hidden.
"Hans hid the football as payback."

c. Hans hat dem Sohn den Fußball versteckt.
Hans has the.DAT son the.ACC football hidden.
"Hans hid the football from the son."

d. Hans hat zur Ahndung dem Sohn den Fußball versteckt.
Hans has as payback the.DAT son the.ACC football hidden.
"Hans hid the football from the son as payback."

Critical to the design of this and previous experiments on German verbal processing is that most clauses are *verb-final*: relatively early in the reading of the sentence, the comprehender obtains sufficient information to infer that the clause will end with a verb (possibly participial, depending on the preceding syntactic context). In the case of the sentences in (3), the key piece of information is the use of the second-position auxiliary verb *hat*, which is a strong sign that the clause is in the present perfect tense, which requires a verb-final participle.[5]

As it stands in (3) above, there is a potential confound in this design: the four versions of the sentence vary in the length of the material preceding the participial verb (we expect locality or expectation effects to appear on this verb). Previous results have indicated that the reading time of a word may be correlated with its position in the sentence (Ferreira & Henderson, 1993; Demberg & Keller, 2008a), which would confound any findings that adding material can speed processing a subsequent verb (and which might indeed confound previous reports of anti-locality effects such as Konieczny, 2000).[6] We address this confound by including additional material preceding the main clause: a subordinate clause with a dative-taking optionally ditransitive verb. This allows us simply to move the dative complement and/or the PP adjunct from the main clause to the subordinate clause to achieve the appropriate configuration for each condition.

The qualitative predictions for this experiment are illustrated in Figure 1. The left panel of this figure graphs the schematic pattern of results that we expect to observe under the expectation-only hypothesis. As we add more phrases to the main clause, processing becomes easier, as the main clause verb becomes more and more expected. Hence (3-a)

[3]Note that an earlier version of DLT (Gibson, 1998) assumes that the discourse referent of the dependent being integrated is also counted for integration cost purposes, i.e., two intervening discourse referents results in an integration cost of three, etc. In the present example, *dem Kunden* would therefore incur an integration cost of three and *aus Freude* would incur an integration cost of one. The difference between the two variants of DLT could be important for verb-final languages, which (unlike English) can be expected to have a larger number of integrations at verbs (two for transitive verbs, three for ditransitive ones, etc.), leading to high integration cost predictions in the 1998 version of the theory.

[4]It is also conceivable that increasing the number of integrations is more important than increasing the length of the integrations; current versions of DLT treat the two as equivalent, but this is not a given.

[5]In the simplified example (3-a), there is a temporary ambiguity such that previous to reading *versteckt*, the auxiliary verb *hat* could be misinterpreted as being a simple-present possessive verb, giving the sentence the meaning "Hans has the football." In our actual experimental materials, however, the contents of the sentence-initial subordinate clause rule out this interpretation nearly categorically; see (4) for a full example item.

[6]We note in passing that this relationship between sentence position and reading time is not universally agreed upon; see, e.g., Vasishth (2003).

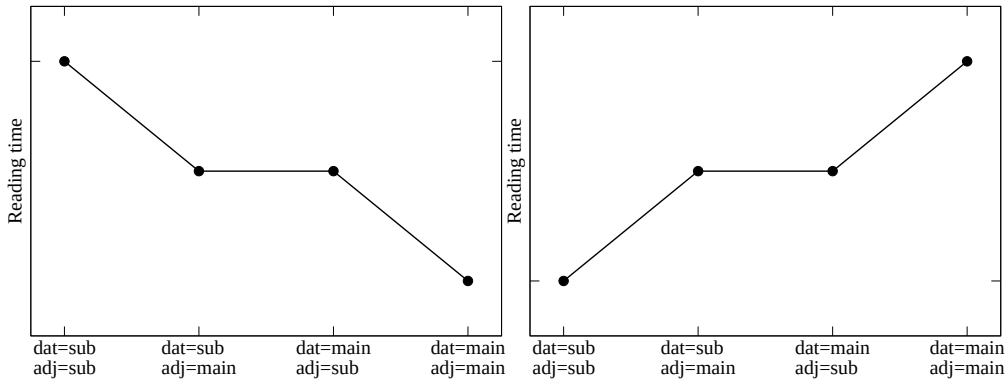

Figure 1: Predictions for Experiments 1 and 2: Left panel: expectation-only hypothesis; right panel: locality-only hypothesis.

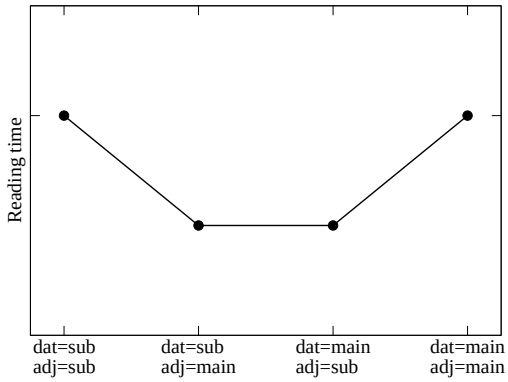

Figure 2: Predictions for Experiments 1 and 2: interaction hypothesis.

(neither dative nor adjunct in the main clause) should be hardest to process, while (3-d) should be easiest (both dative and adjunct in the main clause). (3-b) and (3-c) should be in between (one phrase in the main clause). To the extent that dative NPs and PP adjuncts turn out to have different predictive strength for the clause-final verbs in our materials, however, adding each may have a facilitative main effect of different strength.

The right panel of Figure 1 illustrates the schematic pattern of results expected under a locality-only hypothesis: processing becomes more difficult as we add more phrases to the main clause, because these phrases have to be integrated at the main-clause verb. We would therefore expect (3-a) to be least difficult, (3-d) to be most difficult, and (3-b) and (3-c) to be of intermediate difficulty. To the extent that locality effects from multiple preverbal dependencies are non-additive, the precise predictions regarding reading time in the hardest main/main condition may differ from the schematic depiction of Figure 1.

The results of Konieczny (2000) and Konieczny & Döring (2003) provide some support for the expectation-only hypothesis, as no locality effects were found in their experiment. However, at the beginning of this section, we conjectured that they may have failed to find locality effects because their stimuli did not involve a large enough memory load. If large memory load is required to override expectation effects, then locality effects would be likely to kick in only in the last condition, in which the dative and the adjunct phrase are both in the main clause and have to be integrated. In this case, the experiment should show an interaction of expectation and locality effects, leading to the pattern of results schematically depicted in Figure 2. On this pattern, reading time should decrease if only one of the dative or adjunct appears in the main clause, but should stay the same or increase if both of them are in the main clause.

Corpus Analysis

In order to verify that probabilistic expectations should indeed predict the patterns described above, we conducted a corpus analysis of German main clauses to determine the effects of preverbal dative and adjunct dependents on

	NP.ACC	NP.DAT NP.ACC	PP NP.ACC	PP NP.DAT NP.ACC
$P(w_i = \text{participial verb} \mid w_{1\ldots i-1})$	0.546	0.756	0.504	0.727
Support	1915	78	603	11

Table 1: Conditional probabilities $P(w_i = \text{participial verb} \mid w_{1\ldots i-1})$ for Experiment 1 syntactic configurations, as estimated from tree searches in the combined NEGRA and TIGER corpora.

expectations about the final verb. These expectations can usefully be divided into expectations about (i) *whether* the verb will appear next at any point in online comprehension, and (ii) if the final verb is the next word, *what* verb it may be. In the language of probability theory, these expectations can be described as

(i) $P(w_i = \text{participial verb} \mid w_{1\ldots i-1})$
(ii) $P(w_i \mid w_{1\ldots i-1}, w_i = \text{participial verb})$

The type of corpus data most useful for estimating these probabilities is hand-parsed data such as the NEGRA and TIGER treebanks of German newspaper text (Brants, Skut & Uszkoreit, 1999; Brants, Dipper, Hansen, Lezius & Smith, 2002). Unfortunately, insufficient data are available to easily estimate word-specific probabilities (ii) above for the range of experimental materials we used (our choice of participial verbs was heavily constrained by the requirement that they be optionally ditransitive). Intuitively, however, it seems fairly clear that adding a dative to the main clause should sharpen online expectations in the direction of the participial verbs appearing in our experiment: without the dative, the argument-structure constraints placed on the final verb by its preceding dependents simply limit it to the relatively large set of transitive verbs, whereas adding the preverbal dative NP restricts the final verb to the considerably narrower set of ditransitive verbs. Likewise, it seems fairly clear that adding a PP adjunct should not place as strong a constraint as a dative NP on verb identity, since any verb (subject to the relatively general semantic constraints imposed by the adjuncts we use) can take an adjunct.

We *can*, however, use corpus data to estimate (i), the probability that the next word in a sentence will be a verb, using the syntactic annotation from hand-parsed data. Focusing on the constituency structure of the main clausal constituents, we conducted tree searches in the combined NEGRA and TIGER corpora for syntactic configurations in which a second-position finite auxiliary is followed (not necessarily immediately) in its clause by an accusative NP, with a PP adjunct and/or dative NP possibly preceding the accusative NP in that order; and for extensions of these syntactic configurations in which the accusative NP is immediately followed by a participial verb. These searches (explained in fuller detail in Appendix A) were used to compute relative frequency estimates of probability (i); the results are given in Table 1. Fisher's exact test indicates that conditional verb probability is significantly higher in the dative+accusative-preceding condition than in the accusative-preceding and PP+accusative-preceding conditions ($p < 0.001$ in both cases); few examples were found in the PP+dative+accusative-preceding condition, but the limited data that are available suggest that its behavior is similar to the dative+accusative-preceding condition. These results provide corroboratory evidence for the qualitative nature of the expectation-based predictions given in the previous section.

Method

Participants. Twenty-eight native speakers of German resident in Edinburgh were paid to participate in the experiment.

Stimuli. Thirty-two experimental items were constructed. Each contained a subordinate clause followed by a main clause, both of which were headed by dative-selecting optionally ditransitive verbs. In the subordinate clause, the verb was in simple past tense; in the main clause, the verb was in the present perfect form *hat ... participle*, so that it was clear after reading of the first few words of the main clause that it would end with an obligatory participle. For simplicity, we refer to this final participle as the "clause-final verb" or just the "final verb". We manipulated two factors: the position of a dative NP (subordinate or main clause) and the position of a PP adjunct (subordinate or main clause), and designed each sentence so that all four positionings of the two phrases would result in as natural a sentence as possible. The final verb was followed by a comma (obligatory in German prescriptive grammar) and then by a conjoined participial verb phrase, the beginning of which served as a spillover region (see *Data Analysis*).

In order to ensure high memory load and thus maximize the chances of observing any underlying locality effect at the clause-final verb, we used long dative NPs and long PP adjuncts, each of which introduced two discourse referents. A set of example stimuli is given in (4):

(4) a. Nachdem der Lehrer *zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens* **dem ungezogenen Sohn des fleißigen Hausmeisters** den Strafunterricht verhängte, hat Hans Gerstner den Fußball versteckt, und damit die Sache bereinigt.
After the teacher *as additional payback for multiple wrongdoings* **the.DAT naughty son the.GEN industrious janitor** the.ACC detention.classes imposed, has Hans Gerstner the.ACC football hidden, and thus the affair corrected.
"After the teacher imposed detention classes on the naughty son of the industrious janitor as additional payback for the multiple wrongdoings, Hans Gerstner hid the football, and thus corrected the affair."

b. Nachdem der Lehrer **dem ungezogenen Sohn des fleißigen Hausmeisters** den Strafunterricht verhängte, hat Hans Gerstner *zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens* den Fußball versteckt, und damit die Sache bereinigt.
After the teacher **the.DAT naughty son the.GEN industrious janitor** the.ACC detention.classes imposed, has Hans Gerstner *as additional payback for multiple wrongdoings* the.ACC football hidden, and thus the affair corrected.
"After the teacher imposed detention classes on the naughty son of the industrious janitor, Hans Gerstner hid the football as additional payback for the multiple wrongdoings, and thus corrected the affair."

c. Nachdem der Lehrer *zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens* den Strafunterricht verhängte, hat Hans Gerstner **dem ungezogenen Sohn des fleißigen Hausmeisters** den Fußball versteckt, und damit die Sache bereinigt.
After the teacher *as additional payback for multiple wrongdoings* the.ACC detention.classes imposed, has Hans Gerstner **the.DAT naughty son the.GEN industrious janitor** the.ACC football hidden, and thus the affair corrected.
"After the teacher imposed detention classes as additional payback for the multiple wrongdoings, Hans Gerstner hid the football from the naughty son of the industrious janitor, and thus corrected the affair."

d. Nachdem der Lehrer den Strafunterricht verhängte, hat Hans Gerstner *zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens* **dem ungezogenen Sohn des fleißigen Hausmeisters** den Fußball versteckt, und damit die Sache bereinigt.
After the teacher the.ACC detention.classes imposed, has Hans Gerstner *as additional payback for multiple wrongdoings* **the.DAT naughty son the.GEN industrious janitor** the.ACC football hidden, and thus the affair corrected.
"After the teacher imposed detention classes, Hans Gerstner hid the football from the naughty son of the industrious janitor as additional payback for the multiple wrongdoings, and thus corrected the affair."

The critical region used for analysis was the verb of the main clause (*versteckt* in this example). Note that the memory load at this point is now quite considerable: in DLT terms, for example, in the most extreme condition (example (4-d)) there is an integration cost of four for the PP adjunct (as there are four intervening discourse referents) and two for the dative object, resulting in a total additional cost (beyond that in (4-a)) of six, as opposed to the additional cost of two in (1) from Konieczny & Döring (2003). Additionally, this design rules out a number of confounds that have been present in previous work, ensuring that the sentences are of the same length across all conditions; that exactly the same words precede the critical region (though in different orders) across conditions, so that position in the sentence is identical across conditions; and that the critical region and immediately preceding words are the same across conditions.

General note on statistical data analysis. We used mixed-effects regression models (often called hierarchical or multi-level models Pinheiro & Bates, 2000; Gelman & Hill, 2007; Baayen, Davidson & Bates, 2008) for analysis of all behavioral data in these studies, using R's widely-adopted `lme4` package (Bates, 2011). Mixed effects models have the advantage of allowing the simultaneous consideration of participants and items as random factors in a single analysis, thus avoiding the need for separate F_1 and F_2 and Min F' analyses. Moreover, mixed-effects models are robust in the

face of missing data, a situation that is common in eye-tracking research. For continuous data, including plausibility ratings and reading times, we used linear mixed-effects regression (LMER; Baayen, 2008; Baayen et al., 2008). For categorical data, including Cloze continuation results, and first-pass regression and skip-rate data in the eye-movement studies, we used mixed-effects logistic regression (Jaeger, 2008). This is necessary because the trial-by-trial data for these measures corresponds to a binary response variable, for which standard LMER—which assumes that the response variable is normally distributed around the predicted mean—would be inappropriate (see Baayen 2008, p. 215).

In all analyses we adopt maximal random effects structure, including random slopes for all main effects and interactions. Failing to include random slopes in models when analyzing data with considerable underlying idiosyncratic by-participants or by-items differences can lead to type I errrors in the inferences on fixed effects (see, e.g., Roland, 2009; Barr, Levy, Scheepers & Tily, 2011, for discussion). Thus for a reading-time measure, for exmaple, the formal specification of our model in R's `lme4` package would be `rt ~ dat * adj + (dat * adj|participant) + (dat * adj|item)`. Models are fitted using maximum likelihood (ML; Pinheiro & Bates, 2000) for linear models, Laplace-approximated maximum likelihood for logit models.

The significance of LMER model coefficients is often determined based on *highest posterior density* confidence intervals computed using Markov chain Monte Carlo (MCMC) sampling (see Baayen 2008, p. 270). However, this approach is not available within `lme4` for models with random slopes or for mixed logit models. Instead, we report *p*-values based on normal-approximation interpretation of the *t*-statistic conventionally used in linear regression analysis (this interpretation is suggested by Baayen et al., 2008, and Barr et al. 2011 show that it is only minimally anticonservative for psycholinguistic datasets such as ours, and more conservative than traditional by-subjects and by-items ANOVA), and for mixed logit models we report *p*-values based on the Wald *Z* statistic conventionally used in logistic regression analysis. In all analyses, we center all fixed effects around their means, which minimizes collinearity in analyses of balanced datasets such as ours, and makes main effects fully interpretable even in the presence of interaction terms.

Occasionally we make use of model comparision to assess whether including additional variables signficantly improves model fit. For this we use a likelihood ratio test and evaluate significance against the χ^2 distribution, taking as the degrees of freedom the difference in number of parameters between the two successive models (see Baayen 2008, p. 276). This comparison takes into account the number of parameters of each model and is meant to select the model that gives the most economical and accurate account of the empirical data.

Pretests. It is possible that adding and removing dative complements and PP adjuncts changes the plausibility of the sentences constructed for this experiment. Prior to conducting our reading study, we therefore normed our materials for plausibility in two ways, both to minimize the discrepancy in plausibility across conditions and to use plausibility ratings as a predictor in trial-level data analysis. The first norm, for *global* plausibility, proceeded as follows. As we were primarily interested in reading behavior at the critical verb (*versteckt* in (4) above), we presented versions of our sentences ending in a period immediately after the critical verb, discarding the final *und damit...* phrase but leaving the rest of the sentence untouched. Sixty-one native German speakers rated the plausibility of our experimental items in this form on a scale of 1 (least plausible) to 7 (most plausible). The 128 experimental sentences (32 items in four conditions each) were divided into four lists, such that each list contained exactly one condition of each item, and in any given list, each condition occurred the same number of times (Latin square design). Each list was combined with the same 21 fillers; a separate randomization was generated for each participant. We constructed seven of the fillers to be uncontroversially plausible sentences, seven to involve a highly implausible relationship between a sentence-initial subordinate clause and the main clause, and seven to involve violations of the main verb's argument structure. The pretest was administered over the web using WebExp (Keller, Gunasekharan, Mayo & Corley, 2009).

The mean rating of the plausible fillers was 6.23; of the implausible clause-relationship fillers, 1.96; and of the argument structure violations, 1.67. This wide range indicates that participants made robust distinctions of plausibility for uncontroversially good and bad stimuli. We then used the plausibility ratings for the experimental materials to select the sentences to be used in the subsequent eye-tracking study. An analysis of the mean judgments per item indicated that one item had an untypically low mean rating of 2.39, compared to the global item mean of 4.52. We discarded this item, together with the seven items exhibiting the greatest dative penalty (defined as the difference between the mean scores in the subordinate-clause dative and main-clause dative conditions). This yielded a final set of 24 items with the mean ratings per condition listed in Table 2 (line 1).

Table 2: Mean by-item plausibility ratings per condition for the pretests for Experiment 1; reduced data set with 24 items. The factors are: dative position (dat) and adjunct position (adj), each with the two levels main clause and subordinate clause.

	dat = sub		dat = main	
	adj = sub	adj = main	adj = sub	adj = main
Rating, full sentences	4.54	4.56	4.73	4.70
Raw rating, single clauses	6.06	5.01	5.39	5.04
Length-adjusted rating, single clauses	6.26	5.41	5.79	5.64

Table 3: Plausibilities of fillers in single-clause plausibility norming studies.

Filler type	long	medium	short
Plausible	6.02	6.42	6.38
Slightly implausible	4.15	5.02	4.54
Implausible	3.42	4.12	3.94
Very implausible	1.53	1.64	1.96

An LME analysis on this set showed no significant effect of adjunct position ($t = -0.019$) and no interaction of adjunct and dative position ($t = 0.343$). There was, however, a significant effect of dative position ($t = 2.295$). This effect was small, with a mean dative penalty of .16. We did not expect that reading times would be affected by such a small difference in plausibility—an expectation that was borne out, as will be seen in the *Results* section—and therefore included all 24 of these final items in the eye-tracking study (see Appendix B for the full item set).

Our second norm was for local thematic plausibility of our experimental items, with only the critical clause presented as an independent sentence (e.g., *Hans Gerstner hat den Fußball versteckt* for (4-a)). Since sentence length was not controlled in this norm, we systematically varied sentence length among fillers (including lengths 6, 12, and 18) to determine what effects sentence length on its own might have on plausibility judgments. The fillers (24 in all) were also written to achieve plausibility varying among categories plausible, slightly implausible, implausible, and very implausible. We collected plausibility ratings from 24 native German speakers recruited through Amazon.com's Mechanical Turk in exchange for cash compensation. Mean filler plausibilities are shown in Table 3; LME analyses with maximal random effects structure recovered a significant main effect of plausibility category and a significant linear main effect of length (as measured in number of words). For experimental items, in addition to raw plausibility scores we computed length-adjusted plausibility scores by subtracting the estimated effect of length (as measured from the regression model for fillers) from raw plausibility scores. Both raw and length-adjusted mean plausibility scores are presented in Table 2 (lines 2 and 3 respectively). LME analyses identified a significant main effect of adjunct positioning and a significant interaction between dative and adjunct positioning (all $t > 2$) for both versions of the plausibility scores, and a main effect of dative positioning raw scores ($t = 2.245$) but not for length-adjusted scores ($t = 1.07$). In all cases plausibility scores *dropped* as more material was introduced into the sentence, which runs counter to the anti-locality predictions of expectation-based theories.

Finally, we conducted a Cloze completion study to ensure that the dative argument positioning manipulation successfully affected participants' expectations about verb identity, to complement our corpus analysis. For each item we presented the sentence context up to but not including the critical participle and asked participants to complete the sentence. We collected eight continuations for each item in each condition, with 24 fillers presented alongside to each participant.[7] The second author (a native German speaker) then annotated all continuations, identifying the verbal participle used in each continuation and whether it was a dative-selecting participle (a subjective judgment). Eight completions out of our total 960 did not have a participial verb; we discarded these eight in analyses. Table 4 presents proportions of completions in each condition for which the *first* word in the completion written by the participant

[7]Due to a coding error we failed to include six items to the first 40 participants we recruited for the study. We therefore obtained Cloze completions for these remaining six items from a second round of 40 different participants.

Table 4: Experiment 1 Cloze study results

	dat = sub		dat = main	
	adj = sub	adj = main	adj = sub	adj = main
P(next word is participle used|context)	0.046	0.084	0.134	0.149
P(final word is participle used|context)	0.071	0.121	0.159	0.174
P(final word is dative-selecting participial verb|context)	0.574	0.613	0.965	0.948

was the participle we actually used in the item (true Cloze probability); of completions in which the *final participle, whether it followed immediately or not*, was the participle we actually used in the item; and of completions *where the final participle was a dative-selecting verb*. ANOVA analyses found a significant effect of dative argument positioning on whether the final participle is dative-selecting (both $p \ll 0.001$) and on true Cloze probability (whether the next word is the participle used in the item; by-subjects $p < 0.01$, by-items $p < 0.05$), and a marginal effect on whether the final participle is the one used in the item (both $0.05 < p < 0.1$); no significant effects of adjunct positioning or interactions were recovered. Maximal random-effects logit mixed-effects analysis gave the same results on whether the final participle is dative-selecting (dative: $p \ll 0.001$, adjunct and interaction: $p > 0.45$), but failed to converge for the other two cases.

Procedure. We divided our 96 experimental sentences(24 items in four conditions each) into four lists, such that each list contained exactly one condition of each item, and in any given list, each condition occurred the same number of times (Latin square design). Each list was combined with the same 44 fillers; a separate randomization was generated for each participant. Line breaks were inserted into the items such that the critical region was always in the middle of the third or fourth line, and was always both preceded and followed by at least three words on the same line.

The experiment was run using an Eyelink II head-mounted eye-tracking system, with a sampling rate of 500 Hz. An eye-dominance test was administered for each participant before the experiment began, and only the dominant eye was tracked. A calibration procedure was carried out, and if the calibration was successful, the experiment began. Stimuli were presented with the aid of Eyetrack software developed at the University of Massachusetts at Amherst.[8] Each trial began with a gaze trigger, in the form of a black square that was displayed at the left edge of the screen in a vertically central position. This square occupied the position of the start of the stimulus text, and when the participant fixated the square, it was automatically replaced by the stimulus text. When the participant had finished reading the text, he/she pressed a button on a game pad. Following half of the items, a yes/no question was then displayed on the screen. The participant answered the questions by pressing one of two pre-specified buttons on the game pad. If the automatic gaze contingent stimulus presentation failed on any given trial, the calibration procedure was repeated, and the trial was initiated again. The experimental trials were preceded by four practice trials. The total duration of the experiment was around 35 minutes.

Data Analysis. Vertical drift in the positions of fixations was corrected, using custom software developed at UMass (see footnote 8). An automatic procedure then pooled short contiguous fixations. The procedure incorporated fixations of less than 80 ms into larger fixations within one character, and then deleted any remaining fixations of less than 40 ms. Readers do not extract much information during such short fixations (Rayner & Pollatsek, 1989).

The experimental sentences were divided into four regions. The first region consisted of all words from the start of the sentence up to (but excluding) the main verb. The second, critical, region was the main verb. The third, spill-over, region consisted of the two words following the main verb. The final region comprised the rest of the sentence. The region boundaries for an example sentence are given below:

(5) Nachdem der Lehrer den Strafunterricht verhängte, hat Hans Gerstner zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens dem ungezogenen Sohn des fleißigen Hausmeisters den Fußball/ versteckt,/ und damit/ die Sache bereinigt.

[8]Downloadable from `http://www.psych.umass.edu/eyelab/software/`.

We report data for the following eye-movement measures in the critical and spill-over regions. *First fixation duration* is the duration of the first fixation in a region, provided that there was no earlier fixation on material beyond the region (in which case the first fixation duration is considered zero and the trial is excluded from analysis for first-fixation times). *First pass time* (often called *gaze duration* for single-word regions) consists of the sum of fixation durations beginning with this first fixation in the region until the first saccade out of the region, either to the left or to the right. *Regression path time* is computed as the sum of fixation durations beginning with the first fixation in the region until the first saccade beyond the region. Note that this may include fixations to the left of the region if there is a regression before the reader moves on to the next region. *Total time* consists of the sum of all fixation durations in the region, regardless of when these fixations occur. *Second pass time* consists of the sum of all fixation durations following the first exit of the region (either to right or left). In addition to these reading time measures, we also report the *first pass regressions* measure, which indicates the proportion of trials in which initial first-pass reading is immediately followed by a regressive saccade exiting the region to an earlier part of the sentence.

For all eye-movement measures except second pass time, if on any given trial the region received no fixations, then the data for that trial were not included in the analysis. Moreover, for first-fixation times, first pass time, regression path time and first pass regressions, no trial in which the region is skipped on first-pass reading (i.e., when first-fixation duration is zero) was included in the analysis. We present separate analyses on *skipping rate*, a measure which indicates the proportion of trials in which the region was skipped completely on first-pass reading.

As noted in the Procedure section, we attempted to keep the position of the critical region on the screen relatively constant (always in the middle of the third or fourth line). It is possible, however, that conditions differ in screen positioning in subtle ways, leading to an unwanted confound influencing reading behavior. To evaluate this possiblity, we also report a *launch distance* analysis. The launch distance is the position from which the saccade resulting in the first fixation on the critical region was launched, measured in terms of the number of characters to the left of the critical region. For example a launch distance of eight indicates that the saccade resulting in the first fixation on the critical region was launched from a position eight characters to the left of the beginning of the critical region. Differences in the screen positioning may manifest themselves as differences in launch distance.

Results

Question-answering accuracy

Per-participant accuracy ranged from 67.6% to 89.2%, with a mean of 79.4%. Accuracy did not differ significantly across lists.

Eye movement measures

Table 5 shows the empirical means for the critical and spill-over regions, in the eight eye-movement measures. Table 6 lists the results of the LMER analysis; model coefficients need to be interpreted in the context of our factor coding. Centering of our fixed effects resulted in a value of approximately -0.5 for main-clause positioning of datives and adjuncts, and approximately 0.5 for subordinate-clause positioning.[9] For main effects, positive coefficients thus mean longer reading times or propensity toward regression/skipping when the constituent in question is in the subordinate clause; negative coefficients mean the reverse. Likewise, a positive interaction coefficient indicates that reading time or propensity toward regression/skipping is greater when both the dative and the adjunct phrases are in the same clause.

Table 6 thus indicates that dative position is a significant factor in total time and second pass in the critical region, with shorter reading times when the dative NP is in the main clause. The same effect is present in the spillover region in total and second-pass times. There is a similar effect on first pass times; see later in this section for further details on the statistical significance of this pattern. The dative effect is illustrated by Figure 3, which depicts the model's predicted mean total time in each condition; patterns in the other measures and in the spillover region are similar.

None of the main effects or interactions had a significant effect on launch distance.

The interaction of dative and adjunct position did not reach significance in any of the measures, with the exception of skipping rate in the spill-over region. Here, we find a highly significant interaction of dative and adjunct phrase

[9] In practice the values of the factors will deviate slightly from -0.5 and 0.5 due to slight imbalance from trials with missing data.

Table 5: Empirical means for the eye-movement measures in critical and spill-over regions in Experiment 1 in milliseconds (except for regressions and skipping, which is given as a proportion; and launch distance, which is given in characters). The factors are: dative position (dat) and adjunct position (adj), each with the two levels main clause and subordinate clause.

	dat = sub		dat = main	
	adj = sub	adj = main	adj = sub	adj = main
Critical region				
First fixation	254	277	249	259
First pass	338	354	321	312
Regression path	595	667	649	518
Total time	694	728	637	576
Second pass	318	300	256	216
Regressions	.238	.214	.238	.250
Skipping	.048	.024	.048	.042
Launch site	7.21	7.25	7.53	7.03
Spill-over region				
First fixation	218	211	214	214
First pass	293	313	296	296
Regression path	411	383	369	435
Total time	631	562	538	522
Second pass	293	235	220	202
Regressions	.101	.077	.053	.077
Skipping	.149	.060	.089	.149
Launch site	8.00	7.56	8.38	8.70

position, with a negative coefficient, indicating that there is significantly less skipping if the dative and adjunct phrases are in the same clause. This result merits further discussion, since it does not pattern with any of the other significant results in this experiment. In many of our experimental stimuli, the critical and spillover regions did not lie on the final line of text; close inspection of our stimuli revealed that the number of characters following the spillover region on the same line of text varied considerably across conditions, and patterned similarly with skipping rates: 19.6 characters on average followed the spillover region on the same line in the condition where the dative & adjunct phrases were both in the main clause (the "main/main") condition; 10.2 characters on average in the main/sub condition; 9.6 characters in the sub/main condition, and 16.5 characters in the sub/sub condition. It is possible that readers planned their eye movements on the line of text containing critical and spillover regions such that the final fixation or fixations lay several characters from the end of the text. This would predict that when the number of characters following the spillover region on the same line was small, the skip rate should be lower than otherwise, since the spillover region would fall in an area attracting the last fixations on the line. We tested this hypothesis by entering the number of characters following spillover into a mixed logit regression analysis of skip rate, with random by-subject and by-item slopes. Number of characters following spillover had a highly significant effect in the predicted direction ($\beta = -0.037$ per character, $p < 0.01$) on skip rate. We also entered this covariate together with main effects and interaction of (centered) dative and adjunct position, with random interactions of experimental condition; in this model, number of characters following spillover remained significant ($\beta = -0.071, p = 0.011$) and there was a significant main effect of dative positioning, with more skipping for matrix clause datives ($\beta = 1.26, p < 0.01$) but no effect of adjunct positioning ($p = 0.07$) or interaction ($p = 0.21$). We thus conclude that the dative/adjunct interaction on spillover-region skip rate seen in Table 6 is most likely the product of a confound with the physical positioning of our stimuli on the screen. (The effect we do find, of dative positioning, is theoretically consistent with our other results—matrix-clause datives induce more first-pass skipping.[10])

[10]We nevertheless view this result with a grain of salt; our general impression is that mixed logit analyses in `lme4` with complex random effects structure may be anti-conservative, especially when the overall rates are close to 0 or 1; more systematic exploration along the lines of Barr et al. (2011) has yet to be done, however.

Table 6: Results of linear mixed effects model analysis for Experiment 1. The table lists the coefficients of the best-fitting model with significance level (computed using either MCMC, *t*-test, or *z*-scores, see main text). Note that the factors are centered: "main" is encoded as ≈ -0.5, "sub" as ≈ 0.5. Reading-time coefficients are effect sizes in milliseconds.

	Intercept	dat	adj	dat × adj
Critical region				
First fixation	260.46	11.29	−17.13	−13.64
First pass	332.21	28.54♣	−2.85	−26.47
Regression path	613.73	55.70	41.08	−205.95
Total time	660.62	102.48**	16.28	−96.82
Second pass	272.88	73.26*	28.82	−21.79
Regressions	−1.39	−.11	.01	.26
Skipping	11.62	−11.73	11.29	−27.29*
Launch site	7.25	−.10	.22	−.59
Spill-over region				
First fixation	214.84	0.25	3.48	6.44
First pass	299.03	6.63	−10.94	−20.53
Regression path	400.08	3.63	−15.54	88.96
Total time	557.83	58.98*	29.89	45.54
Second pass	237.85	52.86*	38.18	40.77
Regressions	−2.66	.47	.35	.91
Skipping	3.08	.54	−.04	−2.44**
Launch site	8.28	−.96	.05	.95

$*$: $p < .05$; $**$: $p < .01$; $***$: $p < .001$; ♣: see main text.

To ensure that the other results of this experiment are not being driven by this confound, we also entered number of characters following the spillover region into mixed logit regression analyses with random interactions for all models giving significant results reported in Table 6. Including this covariate had no effect on the qualitative patterns observed. The effect of adjunct on first-fixation times remained marginal; the effect of dative on first-pass times became significant ($t = 1.977$, whereas it was just under significance at $t = 1.941$ without this control variable); the effect of dative on total and second-pass times on both critical and spill-over regions remained significant. Number of characters following the spillover region did show up in these analyses as a significant predictor of total reading times at the critical region and total and second-pass at the spillover region, with shorter reading times in all cases when more characters followed ($\beta = -4.85, -6.38, -3.70$ms per character respectively; all $t > 2$), confirming that the physical positioning of our stimuli on the screen did have some effects on eye movements.

The final remaining possible confound we address is the small global plausibility differential in our items as a function of dative-phrase positioning. (We ignore the local plausibility differential revealed in our second plausibility norming study, since the local plausibility differential went in a direction that was not a confound for our result on dative-phrase positioning effects on eye movement behavior.) Our pretest revealed that structures with a dative NP in the main clause received significantly higher plausibility ratings than structures with a dative in the subordinate clause, even after selecting a more balanced subset (see Table 2). If we assume that processing is easier when the sentence being read is more plausible, then this could explain why we obtained shorter reading times when the dative was positioned in the main clause. We investigated this possibility by fitting a linear mixed effects model that included the factors dative and adjunct position, as well as plausibility, where plausibility was defined as the mean plausibility judgment an item had received in the pretest. The mixed model for total time (formal specification: `rt ~ dat * adj *`

[12] In general, the standard error of the predicted mean response to a predictor vector $\mathbf{X}$ for a new subject/item combination in a linear mixed model is given as $\sqrt{\mathbf{X}^T \Sigma \mathbf{X}}$ where Σ is the covariance matrix of the fixed effects estimate. Because the present design is factorial and balanced, and we have centered our predictors, we can factor the uncertainty about overall reading speed out by omitting the intercept from $\mathbf{X}$ and from Σ—for the dat=0.5,adj=−0.5 condition, for example, we would have $\mathbf{X} = \langle 0.5, -0.5, -0.25 \rangle$. We then use the resulting standard error to compute confidence intervals. This approach is similar to that proposed by Loftus & Masson (1994); the common underlying justification is that "within" designs factor out uncertainty about grand means from analyses of treatment effects, and so confidence intervals should reflect factoring out of this uncertainty.

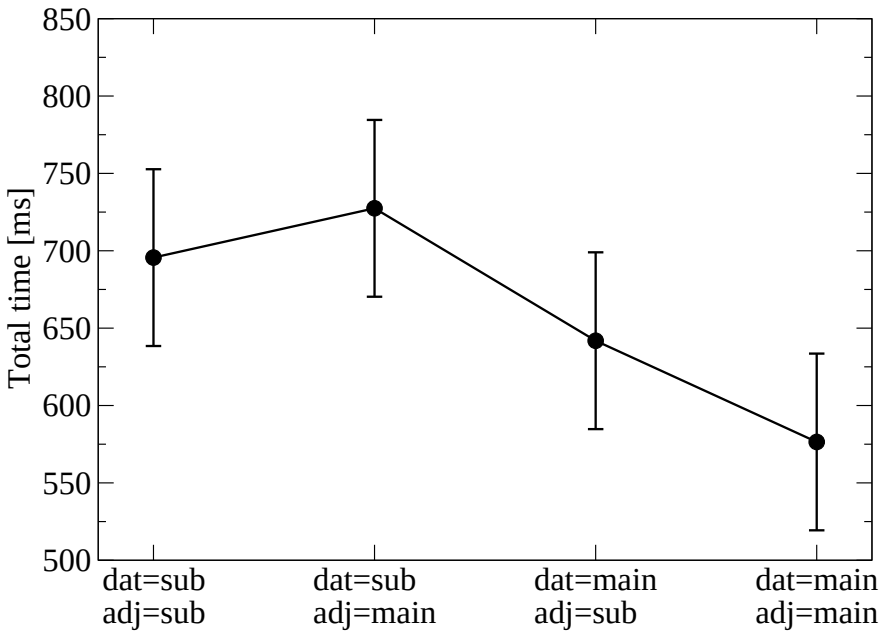

Figure 3: Per-condition means predicted by the mixed model for total time in the critical region for Experiment 1. Error bars show a 95% confidence interval on the predicted mean of each condition, with uncertainty about overall reading speed removed.[12]

`plaus + (dat * adj * plaus|part) + (dat * adj * plaus|item)`) yielded a significant effect of dative position ($\beta = 87.66$ms, $p < .05$), as well as a significant interaction of dative position and plausibility ($\beta = -140.44$, $p < .05$). All other factors and interactions failed to reach significance. We also compared this model to a model without the factor plausibility (formal specification: `rt ~ dat * adj + (dat * adj|part) + (dat * adj|item)`), and failed to find a significant difference in model fit using the log-likelihood criterion ($\chi^2(56) = 36.27$, $p > .1$). In other words, adding plausbility to a model that already contains dative and adjunct placement does not improve its fit with reading time data.[13]

Discussion

We found that the presence of a dative NP in the main clause leads to a decrease in reading time at the main clause's final participial verb. This result, which was observed in first-pass, total, and second pass time, supports the expectation hypothesis: the presence of additional preverbal material makes it possible to generate expectations about the verb, which is then easier to process. The expectation hypothesis is illustrated by Figure 1 (left panel); the results for total time depicted in Figure 3 show the same pattern (though less pronounced, as we found a significant prediction benefit for dative, but not for adjunct phrases).

This result broadly confirms the findings of Konieczny & Döring (2003), even though their study differed from ours in a number of ways: they found an effect in regression path duration (the only measure they report), and their experiment compared the presence of a dative NP with the presence of a genitive NP in the main clause (see (1)), while our study contrasted the presence of the dative NPs with its absence. Unlike the results of Konieczny (2000), our results cannot be attributed to the number of words in the sentence preceding the critical region—which were the same across conditions—or even to the number of words in the same clause preceding the critical region, since the dative and adjunct phrases were of the same length but only introduction of the dative phrase facilitated critical-region reading.

[13]The results for the other measures are consistent with the results for total time. For first fixation, first pass, and second pass, we find that adding plausibility does not improve model fit.

This experiment thus provides support for the expectation hypothesis and no evidence for the locality hypothesis. However, it is possible that our manipulation of memory load may still not have been sufficient to induce appreciable memory-retrieval-based processing difficulty at the final verb, especially considering the ubiquity of verb-final structures in German. Therefore our second experiment introduces an additional factor that might be expected to add to memory load and retrieval difficulty: relativization.

Experiment 2

In Experiment 2, we increased integration cost further by embedding the main clause from the stimuli used in Experiment 1 into a relative clause. This can be illustrated with the following simplified set of materials:

(6) a. Der Mitschüler, der den Fußball versteckt hat, ...
the classmate, who.NOM the football hidden has, ...
"The classmate who hid the football..."

b. Der Mitschüler, der zur Ahndung den Fußball versteckt hat, ...
the classmate, who.NOM as payback the football hidden has, ...
"The classmate who hid the football as payback..."

c. Der Mitschüler, der dem Sohn den Fußball versteckt hat, ...
the classmate, who.NOM the.DAT son the football hidden has, ...
"The classmate who hid the football from the son..."

d. Der Mitschüler, der zur Ahndung dem Sohn den Fußball versteckt hat, ...
the classmate, who.NOM as payback the.DAT son the football hidden has, ...
"The classmate who hid the football from the son as payback..."

Here, the clause that contains the dative or the adjunct NP is a subject relative clause that modifies the head noun *Mitschüler*; the critical region is the verbal complex in the relative clause, *versteckt hat*.

There are several reasons why this use of relativization may increase memory load beyond that in Experiment 1. First, the main-clause structure in Experiment 1 involved a second-position auxiliary verb, whereas the auxiliary verb is final in the relative-clause structure here. If a dependency relation is established between the auxiliary and the subject, having established this relation earlier—as in the main-clause structure—might facilitate subsequent retrieval of the subject when the participle is encountered (e.g., in an interference-based theory such as Lewis et al., 2006, this dependency may have assigned features to the subject that help distinguish it from the other preverbal NPs encountered). Second, relativization induces an *unbounded* dependency, and it is possible that memory retrieval in unbounded dependency construction is especially costly, and/or that the processor devotes more resources to storing an incomplete unbounded dependency than to an incomplete clause-bounded dependency;[14] there would be some logic to such a deployment of working memory resources, since there are fewer guarantees on when an unbounded dependency will be completed.[15] Third, as will be seen in corpus analysis, the most complex syntactic configurations in our experiments we use are less common in relative clauses than in main clauses, and one might expect that storing representations of syntactic configurations imposes lower load on memory when the representations are more frequent.

We thus hypothesize that introducing relativization might increase memory load sufficiently to trigger locality effects, at least in the most extreme condition (6), where the processor now has two additional phrases to integrate (dative and adjunct) on top of dealing with the unbounded dependency of the RC's head noun. If this is the case, then the present experiment should show an interaction of locality and expectation effects on critical-region eye-movement measures, leading to a pattern as in Figure 2.

[14] As an anonymous reviewer points out, the type of the unbounded dependency may also play a role. The stimuli in Experiment 2 include a filler-gap dependency, and resolving it requires consideration of the syntactic configuration, rather than of the individual lexical items involved. It is conceivable that verb-argument dependencies for whom lexical information is more important (argument structure, thematic plausibility) are less prone to locality effects.

[15] One possible interpretation of the results of ERP studies on unbounded dependencies (e.g., Kluender & Kutas, 1993; King & Kutas, 1995; Kaan, Harris, Gibson & Holcomb, 2000; Phillips, Kazanina & Abada, 2005) is that it can be very costly to hold an unbounded dependency in memory for a long time.

	NP.ACC	NP.DAT NP.ACC	PP NP.ACC	PP NP.DAT NP.ACC
$P(w_i = \text{verb} \mid w_{1\ldots i-1})$	0.073	0.103	0.130	0.250
Support	1197	48	269	4

Table 7: Conditional probabilities $P(w_i = \text{participle} \mid w_{1\ldots i-1})$ for Experiment 2 syntactic configurations, as estimated from tree searches in the combined NEGRA and TIGER corpora.

Corpus Analysis

As with the previous experiment, we conducted a corpus analysis to estimate the probability that the next word in a sentence will be a verb, using tree searches on the syntactic annotation from the combined hand-parsed NEGRA and TIGER corpora (full tree-search details given in Appendix A). The results are shown in Table 7. In all cases, the conditional probabilities are much lower than was found in Experiment 1, because previous context does not indicate whether the tense of the relative clause is present perfect. Nevertheless, many of the same tendencies emerge as in Experiment 1: adding a dative NP seems to increase the conditional probability of seeing a participle immediately after the accusative NP (though statistical test results are insignificant due to low support in the dative-present conditions). The main difference is that there is some evidence of facilitation from adding a preceding PP to the accusative-only case ($p < 0.01$ by Fisher's exact test), which was not the case in the main clause covered in Experiment 1. Since the counts of the contexts in which there is a preverbal dative present are so low, it is not clear how seriously we should take the discrepancy between the relative ordering of conditional probabilities of encountering the participial verb next in the dative-only and adjunct-only conditions in Experiment 1 versus Experiment 2.

As for the probability of participle *identity*, we make the same prediction that a preceding dative NP will sharpen expectations more than a preceding PP adjunct, since the argument-structure constraints imposed by the dative NP narrow the space of available verbs more dramatically.

Method

Participants. Twenty-eight participants from the same population as Experiment 2 were recruited for this experiment. None had participated in Experiment 1.

Stimuli. Twenty-four experimental items were constructed (see Appendix B for a full list). These were obtained by modifying the sentences used for Experiment 1 in the following way: the proper name subject of the main clause was replaced by a definite NP modified by a subject-extracted relative clause derived from the main clause in Experiment 1.[16] The spill-over region in Experiment 1 was replaced by a short transitive VP which had as its subject the NP modified by the relative clause. As explained in the introduction to Experiment 2, this manipulation was intended to increase memory load before reaching the final verb, which we hypothesized would make locality effects easier to detect.

As in Experiment 1, the design manipulated two factors: the position of the dative NP (subordinate or relative clause) and the position of a PP modifier (subordinate or relative clause). The presence of the additional material in the relative clause should facilitate the processing of the head verb according to surprisal theory, but make it more difficult according to Dependency Locality Theory. A set of example stimuli is given in (7):

(7) a. Nachdem der Lehrer *zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens* **dem ungezogenen**
After the teacher *as additional payback for multiple wrongdoings* **the.DAT naughty**
Sohn des fleißigen Hausmeisters den Strafunterricht verhängte, hat der
son the.GEN industrious janitor the.ACC detention.classes imposed, has the.NOM
Mitschüler, der den Fußball versteckt hat, die Sache bereinigt.
classmate, who.NOM the.ACC football hidden has, the affair corrected.
"After the teacher imposed detention classes on the naughty son of the industrious janitor as additional payback for the multiple wrongdoings, the classmate who hid the football corrected the affair."

[16] Definite NPs were used in place of proper nouns in this experiment because several subjects in Experiment 1 noted that they occasionally wondered whether the proper-name NP was coreferent with one of the other definite NPs in the sentence.

b. Nachdem der Lehrer **dem ungezogenen Sohn des fleißigen Hausmeisters** den
After the teacher **the.DAT naughty son the.GEN industrious janitor** the.ACC
Strafunterricht verhängte, hat der Mitschüler, der *zur zusätzlichen Ahndung des*
detention.classes imposed, has the.NOM classmate, who.NOM *as additional payback for*
mehrfachen Fehlverhaltens den Fußball versteckt hat, die Sache bereinigt.
multiple wrongdoings the.ACC football hidden has, the affair corrected.
"After the teacher imposed detention classes on the naughty son of the industrious janitor, the classmate who hid the football as additional payback for the multiple wrongdoings corrected the affair."

c. Nachdem der Lehrer *zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens* den
After the teacher *as additional payback for multiple wrongdoings* the.ACC
Strafunterricht verhängte, hat der Mitschüler, der **dem ungezogenen Sohn des**
detention.classes imposed, has the.NOM classmate, who.NOM **the.DAT naughty son the.GEN**
fleißigen Hausmeisters den Fußball versteckt hat, die Sache bereinigt.
industrious janitor the.ACC football hidden has, the affair corrected.
"After the teacher imposed detention classes as additional payback for the multiple wrongdoings, the classmate who hid the football from the naughty son of the industrious janitor, corrected the affair."

d. Nachdem der Lehrer den Strafunterricht verhängte, hat der Mitschüler, der *zur*
After the teacher the.ACC detention.classes imposed, has the.NOM classmate, who.NOM *as*
zusätzlichen Ahndung des mehrfachen Fehlverhaltens **dem ungezogenen Sohn des**
additional payback for multiple wrongdoings **the.DAT naughty son the.GEN**
fleißigen Hausmeisters den Fußball versteckt hat, die Sache bereinigt.
industrious janitor the.ACC football hidden has, the affair corrected.
"After the teacher imposed detention classes, the classmate who hid the football from the naughty son of the industrious janitor as additional payback for the multiple wrongdoings corrected the affair."

As with Experiment 1, this design ensures that the sentences are of the same length across all conditions, ruling out the potential position effects. Due to the effect observed in Experiment 1 of the variable number of characters following the spillover region on the same line, in this experiment we planned line breaks so as to ensure that the critical region and spillover regions (*versteckt hat, die Sache* in the example sentence) occurred on the last line, so that the same number of characters followed these regions in all conditions.

Procedure. The experimental procedure was identical to that used in Experiment 1, except that we attempted to control the horizontal and the vertical position of the critical region controlled more tightly: the critical verb was always on the fourth line of the display, and always the fourth word on that line.

Data Analysis. The experimental sentences were divided into four regions. The first region consisted of all words from the start of the sentence up to (but excluding) the head verb of the relative clause. The second, critical, region was the head participial verb of the relative clause, and the auxiliary that followed it; following this auxiliary there was always a comma marking the end of the relative clause (obligatory in prescriptive German grammar). The third, spill-over, region consisted of the two words following the critical region verb. The final region comprised the rest of the sentence. The region boundaries for an example sentence are given below:

(8) Nachdem der Lehrer den Strafunterricht verhängte, hat der Mitschüler, der zur zusätzlichen Ahndung des mehrfachen Fehlverhaltens dem ungezogenen Sohn des fleißigen Hausmeisters den Fußball/ versteckt hat,/ die Sache/ bereinigt.

The remainder of the data analysis was the same as in Experiment 1. The same eye-movement measures were computed and analyzed using linear mixed effects models as described earlier.

Results

We analyzed the answers participants provided to the comprehension questions. Per-participant accuracy ranged from 62.5% to 96.0%, with a mean of 80.0%. Accuracy did not differ significantly across lists.

Table 8: Empirical means for the eye-movement measures in critical and spill-over regions in Experiment 2 in milliseconds (except for regressions and skipping, which is given as a proportion; and launch distance, which is given in characters). The factors are: dative position (dat) and adjunct position (adj), each with the two levels relative clause and subordinate clause.

	dat = sub		dat = rel	
	adj = sub	adj = rel	adj = sub	adj = rel
Critical region				
First fixation	232	241	254	237
First pass	386	385	377	383
Regression path	565	685	571	642
Total time	793	749	648	742
Second pass	346	268	224	311
Regressions	.143	.244	.170	.161
Skipping	.030	.012	.042	.042
Launch site	7.89	8.28	8.20	7.62
Spill-over region				
First fixation	215	219	208	214
First pass	361	355	362	346
Regression path	720	537	709	778
Total time	798	702	747	846
Second pass	419	342	352	458
Regressions	.101	.071	.065	.071
Skipping	.071	.054	.024	.054
Launch site	7.43	6.53	8.56	7.28

Table 8 shows the empirical means for the critical and spill-over regions, in the seven eye-movement measures. Table 9 lists the results of the LMER analysis; factor coding is -0.5 for constituents positioned in the relative clause, and 0.5 for constituents positioned in the subordinate clause.

As seen in Table 9, there is an interaction of dative and adjunct position significant in second-pass times for the critical and spillover regions and in total times for the spillover region. This interaction bears a positive coefficient, indicating increased reading time occurs when either both the dative and the adjunct phrase are in the subordinate clause, or when both phrases are in the relative clause. The pattern is illustrated in Figure 4, which graphs the means predicted in the LME model for total time at the critical region. (Total time is used for comparability to Figure 3; though the interaction in total time at the critical region is not significant in Experiment 2, the qualitative pattern is the same as second-pass time at the region, for which the interaction is significant.) To further understand this interaction, we conducted comparisons of the effect of dative positioning in each of the adjunct-in-subordinate-clause and adjunct-in-relative-clause positions. These comparisons reveal that positioning the dative in the RC has a significant facilitatory effect when the adjunct is in not in the RC ($t = 2.731$), but has no significant effect when the adjunct is in the RC ($t = -0.999$); the numerical effect is reversed in the two cases. Second-pass and total time at the spillover region show the same pattern, as does total time shows the same pattern, though the latter does not reach statistical significance.

On first-pass regression rates, there is a significant interaction between dative and adjunct such that regression rates are highest either when both or neither phrase is in the matrix clause. There is a significant interaction in the opposite direction in first-pass regressions at the spillover region (but see the following paragraph). Finally, there is a significant interaction on spillover-region skipping, with highest skip rates when both dative and adjunct are in the subordinate clause.

For completeness, we also checked the effect of entering number of characters following the spillover region as a covariate in the models for which significant results are reported in Table 9. All critical and spillover region significant effects remained significant in these analyses, except that the main effect of dative and the dative/adjunct interaction on spillover-region first-pass regressions disappeared. We recovered a significant effect of number of characters following the spillover region on critical-region and spillover-region total times and second-pass times (longer when more characters followed the spillover region), and on skip and first-pass regressions from the spillover region (less skipping

Table 9: Results of linear mixed effects model analysis for Experiment 2. The table lists the coefficients of the best-fitting model with significance level (computed using either MCMC, *t*-test, or *z*-scores, see main text). Note that the factors are centered: "rel" is encoded as ≈ -0.5, "sub" as ≈ 0.5. Reading-time coefficients are effect sizes in milliseconds.

	Intercept	dat	adj	dat × adj
Critical region				
First fixation	241.28	−8.61	3.29	−25.63
First pass	383.01	7.29	−2.69	9.13
Regression path	620.50	15.96	−93.41	−44.68
Total time	733.08	74.46	−22.97	133.29
Second pass	287.94	39.15	−4.06	165.53*
Regressions	−1.83	.02	−.30	−1.14*
Skipping	14.48	18.83	.69	−4.65
Launch site	8.04	.14	.13	−.96
Spill-over region				
First fixation	215.20	6.93	−5.58	−1.06
First pass	352.50	3.27	12.45	−13.69
Regression path	676.73	−129.95	67.78	258.51
Total time	771.84	−40.99	−1.84	187.12*
Second pass	393.15	−24.85	−14.59	183.06*
Regressions	−3.79	1.55*	−.49	3.96***♣
Skipping	7.34	5.83***	−4.96**	−11.96***
Launch site	7.61	−.88	1.03	−.60

$*$: $p < .05$; $**$: $p < .01$; $***$: $p < .001$; ♣: see main text.

and more regressions when more characters followed the spillover region).

Discussion

The results of this experiment matched those of Experiment 1 in the following respect: when the adjunct phrase is outside the RC, including the dative preverbal dependent has a significant facilitatory effect on processing of final verbs, as most clearly indicated in second-pass reading times. When the adjunct phrase is inside the RC, however, we see a different pattern than in Experiment 1: reading times are numerically higher when both phrases occurred in the same clause than when only one phrase was in it (regardless of which phrase). These results can be interpreted as a facilitative prediction effect at the RC verb when the dative NP precedes it as a dependent, but one which is at least partially canceled out when the adjunct phrase also appears as a preverbal dependent in the RC. On this interpretation, the presence of two phrases preceding and dependent on a clause-final verb entails an integration cost high enough to counteract most or all of the prediction benefit obtained from the dative NP.

Recall that the crucial difference between the present experiment and Experiment 1 was the fact that the critical verb was embedded into a relative clause, which we hypothesized might entail increased memory load that could increase difficulty associated with distance-based integration. Experiment 2 supports this hypothesis, and suggests that the processing costs associated with memory load may be *super-additive*: moving the adjunct phrase into the RC does not drive up reading times when the dative NP is in a preceding subordinate clause, but it does when the dative NP is a preverbal dependent inside the RC. To illustrate this, compare Figure 2, which schematically depicts this hypothesis, with Figure 4, which graphs the predicted values for total time in Experiment 2. The patterns are qualitatively similar; the only qualitative difference is that Figure 2 hypothesized expectation-based facilitation from both the adjunct and dative phrases, whereas in both experiments we found expectation-based facilitation only from the dative phrase.

General Discussion

The results in this paper provide evidence for both expectation and locality effects in the processing of verb final clauses in German. We conducted two experiments which tracked participants' eye-movements as they read verb

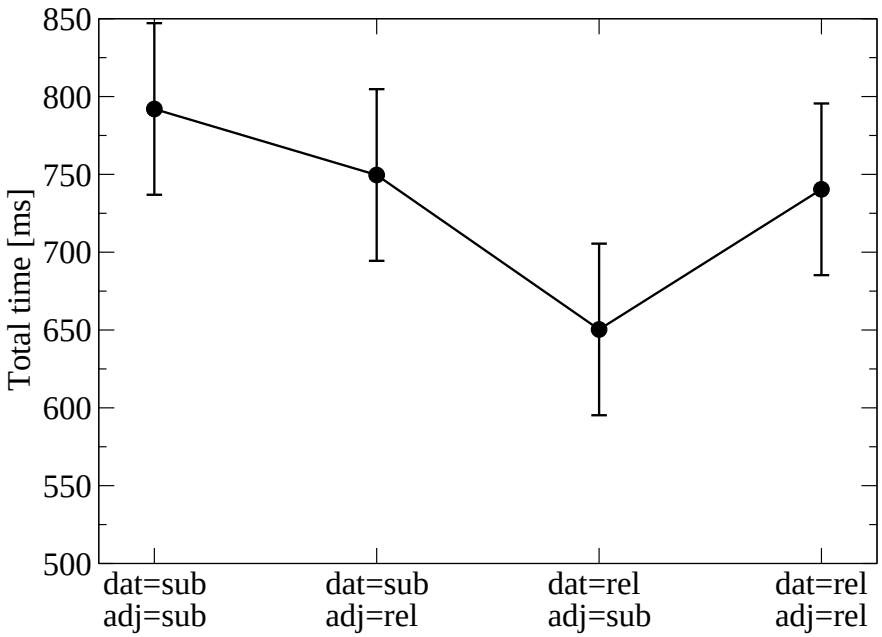

Figure 4: Per-condition means predicted by the mixed model for total time in the critical region for Experiment 2. Error bars show a 95% confidence interval on the predicted mean of each condition, with uncertainty about overall reading speed removed (see Footnote 12).

final dative constructions and found that the presence of a dative noun phrase led to decreased reading time at the corresponding verb, compared to a condition in which there is no preceding dative noun phrase. This can be explained by assuming that the presence the additional preverbal material allows the processor to predict the upcoming verb, which leads to a facilitation effect. In this respect, our experimental results reaffirm the findings of Konieczny (2000) and Konieczny & Döring (2003), while ruling out a number of possible confounds that were present in these earlier experiments.

However, our experiments also showed that locality effects can occur in the same structure: Experiment 2 showed an interaction of adjunct position and dative position, with the verb more difficult to process when both the adjunct and the dative phrase were present than when only one was present. This suggests the presence of a locality effect, i.e., the additional material that needs to be integrated at the verb, leading to a distance-based cost. This effect was only present in Experiment 2, which tested relative clauses, rather than main clauses as in Experiment 1. This suggests that locality effects can override expectation effects under conditions of high memory load, as we hypothesized would be most likely to occur in a relative clause.

Taken together, our results provide evidence for both expectation effects (processing becomes easier with additional material) and locality effects (processing becomes more difficult with additional material). Locality effects have been observed for a range of different constructions in English (Gibson, 1998), while expectation effects (also known as anti-locality effects when the issue is the effect of pre-verbal dependents on verb processing times) have been reported for German, Japanese, and Hindi (Konieczny, 2000; Konieczny & Döring, 2003; Nakatani & Gibson, 2008; Vasishth & Lewis, 2006), and more recently also for English (Demberg & Keller, 2008a; Jaeger et al., 2008). However, ours is the first demonstration to our knowledge that both expectation and locality effects can occur in the same structure in the same language, and that the two effects interact with each other. More specifically, we found that expectation effects dominate locality effects, insofar as Experiment 1 did not demonstrate any locality effects. Locality was only able to counteract expectation in cases of extremely high memory load, as in the longest relative clauses used in Experiment 2.

In our experiments, most of the empirical effects relevant to these issues emerge most clearly in later eye movement measures, notably second-pass and total reading times. This contrasts to some degree with the results of Konieczny & Döring (2003), who found expectation-based facilitation in regression-path durations (though note that in Experiment 1 we did find significant facilitation from dative preverbal dependents once physical positioning on the screen was

introduced as a control variable). It is not entirely clear to us why the crucial effects in our experiments mostly showed up during late measures, when, for example, it is well established that expectation-based facilitation occurs often in early measures (e.g., Ehrlich & Rayner 1981). One possibility is that due to the unusual length and complexity of our sentences for a psycholinguistic study—necessitated by our design, in particular its elimination of sentence-position confounds—our participants may have adopted a strategy of re-reading entire sentences a second time through rather than attempting to do comprehensive processing the first time around, so that that much of the processing differential in and just after the critical region is picked up in the second pass rather than in the first pass.

None of the classical theories of online syntactic processing complexity is straightforwardly compatible with our results. The existence of expectation effects is inconsistent with Gibson's 1998 Dependency Locality Theory, which predicts that additional material that needs to integrated with a head increases processing effort at the head, which is the opposite of what we found in the case of dative phrases in our experiments. Surprisal theory (Hale, 2001) does predict the expectation effects we found, as demonstrated by Levy's 2008a surprisal-based model of the closely related findings of Konieczny (2000) and Konieczny & Döring (2003). However, surprisal alone is not sufficient to explain why under some circumstances what seem to be locality effects can partially override expectation-based facilitation in the same structure, as we saw in Experiment 2. These results call for the development of new models of processing difficulty in online sentence comprehension which integrate the insights of surprisal and locality theories. Similar conclusions have been reached by other researchers, including Grodner & Gibson (2005), who noted the apparent conflict between locality-based reading-time patterns they observed in English relative clauses and expectation-based patterns found by Konieczny (2000), Vasishth (2002), and Nakatani & Gibson (2003; now published as Nakatani & Gibson, 2010); by Vasishth & Lewis (2006), who explicitly investigated the differences in reading-time patterns in English versus Hindi relative clauses; by Demberg & Keller (2008a), who show that both integration cost and surprisal effects can be found in an eye-tracking corpus; by Patil, Vasishth & Kliegl (2009) who find both memory retrieval cost and surprisal effects, again in an eye-tracking corpus; and by Vasishth & Drenhaus (2011), who have recently found evidence for locality effects on German verbal processing. The present experiments make an important additional contribution to this literature, however, as the first simultaneous demonstration of locality and anti-locality effects in controlled experiments using a single grammatical construction in a single language—verb-final clauses in German—achieved by varying both expectations and memory load through the number and type of preverbal dependents and the presence of relativization.

We close with brief speculation on the reasons why we see the particular tradeoffs between expectations and locality that we do, and prospects for developing precise, implemented models that can account for both locality and anti-locality effects of the sorts found here. As noted in the Introduction, memory-limitation effects of the type predicted by DLT and SBI theories have been far more elusive for verb-final languages such as German, Japanese, and Hindi than for verb-medial languages such as English, French and Russian. An intuitive possible explanation for this is that native speakers of verb-final languages are simply more practiced and therefore more skilled at comprehending non-local syntactic configurations. It has recently been shown that average total-sentence dependency lengths are considerably longer in German than in English (Gildea & Temperley, 2010; Park & Levy, 2009). Corroborative evidence can also be found in Vasishth, Suckow, Lewis & Kern (2010), who find that German speakers are better than English speakers at tracking multiple incomplete noun-verb dependencies induced by multiple center-embedding. Putting all these results together, it appears that theories of syntactic complexity may need to posit memory costs which are a function of a speaker's linguistic experience rather than fixed and universal. Crucially, however, our results add to what has thus far been only a small amount of evidence (Jaeger et al., 2008; Vasishth & Drenhaus, 2011) that processing patterns at verbs may not show a categorical either/or pattern in any language, but rather a combination of expectation- and memory-based patterns. A natural inference is that this may be universally true, and that the difference in the relative dominance of locality versus expectation for any particular language may be a matter of degree (presumably derivable from the language's grammatical properties).

Regarding the development of more precise models of these phenomena, it is useful here to briefly consider a distinction made in much of cognitive science (Marr, 1982; Anderson, 1990) between models developed at the *algorithmic* level—those which focus on input and output representations and the steps taken to convert inputs to outputs—and those developed at the *computational* level—those which focus on the ultimate goals of computation, constraints on the computing system, and what properties might hold of a well-designed system for achieving the goals subject to the constraints. At the algorithmic level, at least two theories are available which could accommodate our results. The first

is Lewis & Vasishth's 2005 ACT-R based model of sentence processing as skilled memory retrieval (see also Vasishth & Lewis, 2006; Lewis et al., 2006). In this model, processing of a clause-final verb requires retrieval and integration of its preceding dependents, a process which is subject to interference from preceding material. Greater amounts and complexity of preceding material yield greater interference effects, which could explain the difficulty seen in the dative-RC/adjunct-RC condition of Experiment 2. At the same time, however, preceding dependents boost the activation of the upcoming verb and thus facilitate its retrieval; under the appropriate conditions, then, the appropriate quantity and type of preceding material can facilitate final-verb processing, as seen in both our experiments.

The second algorithmic theory readily available to account for our findings is the Psycholinguistically Motivated Tree-Adjoining Grammar (PLTAG) model of Demberg & Keller (2008b, 2009), a model of incremental syntactic analysis through tree-fragment combination which also includes an explicit syntactic-prediction component. This component generates predictions of upcoming syntactic material if this material is required to ensure word-by-word incrementality while maintaining a fully connected tree. In PLTAG, predictions can be read directly off the PLTAG tree, and surprisal scores can be computed based on a probability distribution over partial PLTAG trees defined by Demberg & Keller (2009). In addition to surprisal-based expectation effects, however, the PLTAG model includes a verification component, which closely mirrors DLT's memory cost component. Whenever a prediction is discharged, a penalty is incurred that is based on the distance between the first generation of a prediction and its verification. Demberg & Keller (2009) show that their model can account for locality effects in subject vs. object relative clauses as well as for expectation effects in coordinate structures.

At the computational level, it is possible that an analysis of our findings could be available within the uncertain-input rational comprehension model of Levy (2008b). This model does not contain an explicit memory-retrieval or integration cost for the online processing of verbs (or, for that matter, any other syntactic category). However, both the representations of current sensory input and memory traces from previous sensory input are taken in this model to be noisy and uncertain. When comprehending simple and relatively canonical types of sentences, these representations are corrected and sharpened by prior grammatical and world knowledge. When comprehending complex sentences with infrequent grammatical structures, however, prior knowledge is less effective in reinforcing these representations. In such situations, the representation of preceding context may not be sharp enough to support the expectations about upcoming input that would be computed by an ideal comprehender with perfect perception and memory. We might thus see a failure to take advantage of what should ideally be accurate expectations in the comprehension of the most complex sentences, as in Experiment 2.

In all three cases, substantial work remains to be done to demonstrate whether these speculative explanations can in fact account for the data reported here within the existing theoretical frameworks. Since the reconciliation of locality and anti-locality effects has become an area of substantial interest for the sentence-processing community, we plan to conduct additional modeling work in these directions, and hope that this work, together with new empirical studies, can qualitatively improve our understanding of the interplay between memory and probabilistic knowledge in language processing.

Acknowledgments

Portions of this work have benefited from presentation at the 2007 and 2008 CUNY Sentence Processing Conferences, and from feedback at colloquia given at the University of Edinburgh and the University of Rochester. We are grateful to Patrick Sturt and Keith Rayner for advice regarding data analysis and interpretation. Any errors are our own. RPL was supported in this work by ESRC Postdoctoral Fellowship award PTA-026-27-0944, a travel grant from the Royal Society of Edinburgh, NSF grant 0953870, and NIH grant 1R01HD065829. FK was supported by EPSRC research grant EP/C546830/1: Prediction in Human Parsing.

References

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

Anderson, J. R. (1990). *The Adaptive Character of Human Thought*. Lawrence Erlbaum.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2011). Random effects structure in mixed-effects models: Keep it maximal. Under review.

Bates, D. (2011). Linear mixed model implementation in `lme4`. Manuscript, University of Wisconsin, 4 October 2011.

Brants, S., Dipper, S., Hansen, S., Lezius, W., & Smith, G. (2002). The TIGER treebank. In *Proceedings of the Workshop on Treebanks and Linguistic Theories*.

Brants, T., Skut, W., & Uszkoreit, H. (1999). Syntactic annotation of a German newspaper corpus. In *Proceedings of the ATALA Treebank Workshop* (pp. 69–76).

Demberg, V., & Keller, F. (2008a). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 101*, 193–210.

Demberg, V., & Keller, F. (2008b). A psycholinguistically motivated version of TAG. In *Proceedings of the 9th International Workshop on Tree Adjoining Grammars and Related Formalisms*. Tübingen.

Demberg, V., & Keller, F. (2009). A computational model of prediction in human parsing: Unifying locality and surprisal effects. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Amsterdam: Cognitive Science Society.

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

Ferreira, F., & Henderson, J. M. (1993). Reading processes during syntactic analysis and reanalysis. *Canadian Journal of Experimental Psychology*, *16*, 555–568.

Garnsey, S. M., Perlmutter, N. J., Meyers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, Language, Brain* (pp. 95–126). MIT Press.

Gildea, D., & Temperley, D. (2010). Do grammars minimize dependency length? *Cognitive Science*, *34*, 286–310.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 1411–1423.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97–114.

Gordon, P. C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 1304–1321.

Grodner, D., & Gibson, E. (2005). Some consequences of the serial nature of linguistic input. *Cognitive Science*, *29*, 261–290.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the 2nd Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 159–166). Pittsburgh, PA: Association for Computational Linguistics volume 2.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition, 90*, 3–27.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

Jaeger, T. F., Fedorenko, E., Hofmeister, P., & Gibson, E. (2008). Expectation-based syntactic processing: Anti-locality outside of head-final languages. Oral presentation at CUNY 2008.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of integration difficulty. *Language and Cognitive Processes, 15*, 159–201.

Keller, F., Gunasekharan, S., Mayo, N., & Corley, M. (2009). Timing accuracy of web experiments: A case study using the WebExp software package. *Behavior Research Methods, 41*, 1–12.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7*, 376–395.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*, 196–214.

Konieczny, L. (1996). *Human sentence processing: a semantics-oriented parsing approach*. Ph.D. thesis Universität Freiburg.

Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research, 29*, 627–645.

Konieczny, L., & Döring, P. (2003). Anticipation of clause-final heads: Evidence from eyetracking and SRNs. In *Proceedings of the 4th International Conference on Cognitive Science and the 7th Conference of the Australasian Society for Cognitive Science* (pp. 330–335). University of New South Wales, Sydney.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161–163.

Lau, E., Stroud, C., Plesch, S., & Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language, 98*, 74–88.

Levy, R. (2005). *Probabilistic Models of Word Order and Syntactic Discontinuity*. Ph.D. thesis Stanford University.

Levy, R. (2008a). Expectation-based syntactic comprehension. *Cognition, 106*, 1126–1177.

Levy, R. (2008b). A noisy-channel model of rational human sentence comprehension under uncertain input. In *Proceedings of the 13th Conference on Empirical Methods in Natural Language Processing* (pp. 234–243).

Levy, R., & Andrew, G. (2006). Tregex and Tsurgeon: tools for querying and manipulating tree data structures. In *Proceedings of the 2006 conference on Language Resources and Evaluation*.

Levy, R., Fedorenko, E., Breen, M., & Gibson, T. (2012). The processing of extraposed structures in English. *Cognition, 122*, 12–36.

Levy, R., Fedorenko, E., & Gibson, E. (2011). The syntactic complexity of Russian relative clauses. Under review.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*, 1–45.

Lewis, R. L., Vasishth, S., & Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science*, *10*.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Mitchell, D. C. (1984). An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In D. Kieras, & M. A. Just (Eds.), *New methods in reading comprehension*. Hillsdale, NJ: Earlbaum.

Nakatani, K., & Gibson, E. (2003). An on-line study of Japanese nesting complexity. Presented at the 2003 annual CUNY Sentence Processing Conerence.

Nakatani, K., & Gibson, E. (2008). Distinguishing theories of syntactic expectation cost in sentence comprehension: Evidence from Japanese. *Linguistics*, *46*, 63–86.

Nakatani, K., & Gibson, E. (2010). An on-line study of Japanese nesting complexity. *Cognitive Science*, *34*, 94–112.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, *11*, 283–334.

Park, Y. A., & Levy, R. (2009). Minimal-length linearizations for mildly context-sensitive dependency trees. In *Proceedings of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL-HLT) conference*.

Patil, U., Vasishth, S., & Kliegl, R. (2009). Compound effect of probabilistic disambiguation and memory retrievals on sentence processing: Evidence from an eye-tracking corpus. In A. Howes, D. Peebles, & R. Cooper (Eds.), *Proceedings of 9th International Conference on Cognitive Modeling*. Manchester.

Phillips, C., Kazanina, N., & Abada, S. H. (2005). Erp effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*, 407–428.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading*. Englewood Cliffs, NJ: Prentice Hall.

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

Roland, D. (2009). Relative clauses remodeled: The problem with mixed-effect models. Poster presentation at the 2009 CUNY Sentence Processing Conference.

Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from *Either . . . Or*. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 425–436.

Staub, A., Clifton, C., & Frazier, L. (2006). Heavy NP shift is the parser's last resort: Evidence from eye movement. *Journal of Memory and Language*, .

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

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, *35*, 454–475.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

Vasishth, S. (2002). *Working memory in sentence comprehension: Processing Hindi center embeddings*. Ph.D. thesis Ohio State University.

Vasishth, S. (2003). Quantifying processing difficulty in human sentence parsing: The role of decay, activation, and similarity-based interference. In *Proceedings of EuroCogSci*.

Vasishth, S., & Drenhaus, H. (2011). Locality effects in german. *Dialogue and Discourse, In press*.

Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language, 82*, 767–794.

Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language & Cognitive Processes, 25*, 533–567.

Appendix A. Tree Search Patterns

All searches were conducted using the Tregex tree-search software (Levy & Andrew, 2006), on a version of the combined NEGRA and TIGER corpus transformed automatically to context-free format (Levy, 2005).

Experiment 1: main-clause searches

The term `*aux*` is used here as a place-holder for the expression `/^(hat|haben|hatte|hatten|ist|sind|war|w.ren)$/`, which picks out finite third-person forms of the auxiliary verbs *haben* and *sein*.

Configuration	Pattern	Count
NP.ACC	`VAFIN-HD < *aux* .. NP-OA=obj $++ (@VP <- (__ ,, =obj))`	1915
...immediately followed by verb	`VAFIN-HD < *aux* $++ (@VP < (@VVPP , NP-OA))`	1046
NP.DAT preceding NP.ACC	`VAFIN-HD < *aux* .. (NP-DA .. NP-OA=obj) $++ (@VP <- (__ ,, =obj))`	78
...immediately followed by verb	`VAFIN-HD < *aux* .. (NP-DA .. NP-OA=obj) $++ (@VP < (@VVPP , =obj))`	59
PP adjunct preceding NP.ACC	`VAFIN-HD < *aux* .. (PP-MO.. NP-OA=obj) $++ (@VP <- (__ ,, =obj))`	603
...immediately followed by verb	`VAFIN-HD < *aux* .. (PP-MO.. NP-OA=obj) $++ (@VP <- (@VVPP , =obj))`	304
PP adjunct preceding NP.DAT preceding NP.ACC	`VAFIN-HD < *aux* .. (PP-MO .. (NP-DA .. NP-OA=obj)) $++ (@VP <- (__ ,, =obj))`	11
...immediately followed by verb	`VAFIN-HD < *aux* .. (PP-MO .. (NP-DA .. NP-OA=obj)) $++ (@VP <- (@VVPP , =obj))`	8

Experiment 2: subject-extracted relative clause searches

These searches were made slightly more complex by the fact that the finite verb in German relative clauses is itself clause-final, so that preverbal information in the clause does not clearly indicate that there will be a participial verb at all. As a result, the ratio between conditioning counts and verb-outcome counts is much higher in these searches than for the main-clause searches.

Configuration	Pattern	Count
NP.ACC	`S-RC < (PRELS-SB .. NP-OA=obj) < (/^V.FIN/ ,, =obj)`	1197
...immediately followed by verb	`S-RC < (PRELS-SB .. NP-OA=obj) < (/^V.FIN/ $- (@VP < (@VVPP , =obj)))`	87
NP.DAT preceding NP.ACC	`S-RC < (PRELS-SB .. (NP-DA .. NP-OA=obj)) < (/^V.FIN/ ,, =obj)`	48
...immediately followed by verb	`S-RC < (PRELS-SB .. (NP-DA .. NP-OA=obj)) < (/^V.FIN/ $- (@VP < (@VVPP , =obj)))`	5
PP adjunct preceding NP.ACC	`S-RC < (PRELS-SB .. (PP-MO.. NP-OA=obj)) < (/^V.FIN/ ,, =obj)`	269
...immediately followed by verb	`S-RC < (PRELS-SB .. (PP-MO.. NP-OA=obj)) < (/^V.FIN/ $- (@VP < (@VVPP , =obj)))`	35
PP adjunct preceding NP.DAT preceding NP.ACC	`S-RC < (PRELS-SB .. (PP-MO .. (NP-DA .. NP-OA=obj))) < (/^V.FIN/ ,, =obj)`	4
...immediately followed by verb	`S-RC < (PRELS-SB .. (PP-MO .. (NP-DA .. NP-OA=obj))) < (/^V.FIN/ $- (@VP < (@VVPP , =obj)))`	1

Appendix B. Experimental Materials

Experiment 1

We give the experimental stimuli in the condition dat = main, adj = main. The other conditions can be constructed by moving the dative phrase and/or the adjective phrase into the subordinate clause, as illustrated by example (4) in the main text.

(1) Weil der Verkäufer den Mangel verheimlichte, hat Peter Mühlberger ohne langes Bedenken des
Because the seller the defect hid has Peter Mühlberger without long consideration of-the
problematischen Sachverhalts dem arroganten Händler des vornehmen Autohauses den Höchstpreis geboten, und
problematic fact the arrogant dealer of-the posh car-dealership the maximum-price offered and
so einen Verlust gemacht.
thus a loss made

(2) Nachdem der Lehrer den Strafunterricht verhängte, hat Hans Gerstner zur zusätzlichen Ahndung des mehrfachen
after the teacher the detention-classes imposed has Hand Gerstner as additional payback of-the multiple
Fehlverhaltens dem ungezogenen Sohn des fleißigen Hausmeisters den Fußball versteckt, und damit die Sache
wrongdoings the naughty son of-the industrious janitor the football hidden and thus the affair
bereinigt.
corrected

(3) Weil der Kollege den Bestechungsvorwurf machte, hat Fritz Lorzig nach intensiver Diskussion des brisanten Falls
because the colleague the corruption-allegation made has Fritz Lorzig after intensive discussion of-the delicate case
dem fleißigen Mitarbeiter des strengen Chefs den Aktenkoffer entwendet, und damit eine Straftat begangen.
the industrious worker of-the strict boss the briefcase stolen and thus a crime committed

(4) Weil der Stadtrat den Antrag ablehnte, hat Jens Hartmann nach stundenlanger Erörterung des langen Berichts
because the councilor the application rejected has Jens Hartmann after many-hour discussion of-the long report
dem schmierigen Politiker des kleinen Ortes den Finanzplan präsentiert, und so die Krise entschärft.
the sleazy politician of-the small town the financial-plan presented and therefore the crisis defused

(5) Weil der Manager den Vorschlag zurückwies, hat Wolfgang Meier nach sofortiger Angabe des überzeugenden
because the manager the proposal rejected has Wolfgang Meier after immediate mention of-a convincing
Grunds dem senilen Vorstand des unrentablen Unternehmens den Aktienanteil entzogen, und damit die
reason the senile executive of-the unprofitable company the share-holding withdrawn and therefore the
Pleite verursacht.
bankruptcy caused

(6) Obwohl der Komplize den Zeugendienst verweigerte, hat Karl Riester nach stundenlangem Abstreiten des fraglichen
although the accessory the witness-service refused has Karl Riester after many-hour denial of-the contested
Sachverhalts dem wütenden Polizisten des kleinen Orts den Mittäter genannt, und damit den Fall abgeschlossen.
fact the angry policeman of-the small town the accomplice named, and thus the case closed

(7) Als der Vertreter den Vertriebsplan vorstellte, hat Thomas Schmidt nach erfolgreicher Widerlegung des
As the representative the distribution-plan presented has Thomas Schmidt after successful refutation of-the
unbegründeten Einwands dem ungeduldigen Vorsitzenden des großen Konzerns den Vertrag gesichert, und so ein
unfounded objection the impatient head of-the big trust the contract secured and thus a
Riesengeschäft gemacht.
huge-business made

(8) Weil der Einkäufer den Preis erhöhte, hat Frank Sailer nach erneuter Ausweitung des großzügigen Angebots dem
Because the buyer the price increased has Frank Sailer after renewed expansion of-the generous offer the
kränkelnden Betrieb des alternden Unternehmers den Auftrag versprochen, und damit den Konkurs abgewendet.
sickly company of-the aging entrepreneur the order promised and therefore the bankruptcy prevented

(9) Nachdem der Beamte den Schadensfall bescheinigte, hat Sabine Berger trotz vereinzelter Bedenken des weitläufigen
After the official the damage certified, has Sabine Berger despite isolated worried of-the large
Familienkreises dem effizienten Versorgungsamt des kleinen Kurorts den Rentenantrag gestellt, und dann den
family-circle the efficient benefit-office of-the small resort the pension-application submitted, and then the
Lebensabend genossen
retirement enjoyed

(10) Nachdem der Konzern den Filmvertrags anbot, hat Karin Kowalsky trotz heftigen Protestes des verärgerten
After the trust the move-contract offered has Karin Kowalsky despite intense protests of-the annoyed
Produzenten dem berühmten Regisseur des erstklassigen Werks den Schauspieler vorgestellt, und so die Produktion
producer the famous director of-the first-class work the actor introduced and thus the production
ermöglicht.
enabled

(11) Nachdem der Manager den Bonus auszahlte, hat Heike Schneider wegen guter Leistung des restrukturierten
after the manager the bonus paid-out has Heike Schneider because-of good performance of-the restructured
Außendiensts dem verdienten Mitarbeiter des profitablen Betriebs den Vertrag verlängert, und damit
sales-force the acclaimed worker of-the profitable company the contract extended and thus
Weitsicht bewiesen.
forward-planning shown

(12) Nachdem der Koch den Truthahn zubereitete, hat Simone Küfer unter genauer Beachtung des detaillierten Reglements
After the cook the turkey prepared has Simone Küfer under exact following of-the detailed regulations
dem treuen Stammgast des französischen Restaurants den Hauptgang serviert, und so den Kritiker beschwichtigt.
the faithful customer of-the French restaurant the main-course served and thus the critic placated

(13) Nachdem der Trainer den Übungsplan entwickelte, hat Birgit Messner trotz negativer Berichterstattung des
after the coach the lesson-plan developed has Birgit Messner despite negative reporting of-the
dummen Journalisten dem ehrgeizigen Sportler des aufgestiegenen Vereins den Erfolg vorhergesagt, und damit
stupid journalist the ambitious sportsman of-the promoted club the success predicted and therefore
Recht behalten.
right remained

(14) Weil der Landtag den Zuschuss bewilligte, hat Franz Ritter nach langwierigem Aushandeln des schwierigen
because the parliament the subsidy approved has Franz Ritter after long negotiations of-the difficult
Vertrags dem kleinen Betrieb des ehemaligen Zonenrandgebiets den Auftrag eingebracht, und damit Arbeitsplätze
contract the small company of-the former border-area the order earned and thus jobs
gesichert.
secured

(15) Weil der Gefängnisdirektor den Hafturlaub genehmigte, hat Stefan Bauer nach vorschriftsmäßigem Befragen
because the prison-director the prison-holiday approved has Stefan Bauer after rule-compliant consultation
des direkten Vorgesetzten dem harmlosen Häftling des kleinen Gefängnisses den Ausgang erlaubt, und so den
of-the direct superior the harmless inmate of-the small prison the leave granted and thus the
Besuch ermöglicht.
visit enabled

(16) Nachdem der Geizhals den Schuppen abriss, hat David Weinberger ohne große Beachtung des kalten Wetters
after the scrooge the shed demolished has David Weinberger without much regard to-the cold weather
dem frierenden Stadtstreicher des armen Vororts den Sommermantel geschenkt, und damit Unruhe gestiftet.
the freezing hobo of-the poor suburb the summer-coat given and thus unrest caused

(17) Weil der Vater den Ratschlag gab, hat Klaus Berger nach nochmaligem Überdenken des schwierigen Problems
because the father the advice gave has Klaus Berger after repeated contemplation of-the difficult problem
dem reichen Onkel des entfernten Verwandten den Bittbrief geschickt, und damit die Notlage beseitigt.
the rich Uncle of-the distant relative the request-letter sent and therefore the distress removed

(18) Nachdem der Junge den Hilfsdienst versprach, hat Martin Kaufmann ohne langes Erwägen des möglichen
After the boy the help promised has Martin Kaufmann without long deliberation of-the possible
Trinkgelds dem alten Rentner des baufälligen Nachbarhauses den Koffer getragen, und so einen guten
tip the old pensioner of-the dilapidated neighbor-house the suitcase carried and thus a good
Eindruck gemacht.
impression made

(19) Obgleich der Manager den Zweifel anmeldete, hat Lisa Friedländer nach genauer Überprüfung des detaillierten
Even-though the manager the doubt raised has Lisa Friedländer after exact verification of-the detailed
Angebots dem dynamischen Vertreter des kleinen Unternehmens den Auftrag vergeben, und so die Firma
offer the dynamic representative of-the small business the order given and so the company
ruiniert.
ruined

(20) Nachdem der Polizist den Durchsuchungsbefehl vorlegte, hat Heinz Mischnik trotz sofortigen Einspruches des
After the policeman the search-warrant presented has Heinz Mischnik despite immediate objection of-the
gewitzten Anwalts dem beharrlichen Fahnder des örtlichen Finanzamts den Aktenschrank aufgemacht, und dann
clever lawyer the insistent investigator of-the local tax-office the file opened and then
die Beweismittel übergeben.
the evidence handed-over

(21) Weil der Pächter den Beistand leistete, hat Paul Schnitzer zur sofortigen Behebung des offensichtlichen
because the tenant the assistance rendered has Paul Schnitzer for immediate remedy of-the obvious
Notstands dem frierenden Bewohner des eingestürzten Hochhauses den Schuppen aufgeschlossen, und so die Not
emergency the freezing occupant of-the collapsed tower-block the shed unlocked and so the distress
gelindert.
reduced

(22) Weil der Großvater den Aktienanteil verkaufte, hat Johannes Probst trotz erheblicher Bedenken des
because the grandfather the shares sold has Johannes Probst despite serious doubts of-the
gewissenhaften Steuerberaters dem gierigen Angestellten des großen Finanzhauses den Anteilschein vorgelegt,
meticulous tax-accountant of-the greedy employee of-the big financial-company the coupon presented
und damit das Vermögen gefährdet.
and thus the fortune endangered

(23) Nachdem der Passant den Sachverhalts schilderte, hat Markus Kaiser trotz anfänglichem Bestreiten des
after the passer-by the fact described has Markus Kaiser despite initial denial of-the
peinlichen Vorfalls dem jun gen Beamten des herbeigerufenen Streifenwagens den Ausweis
embarrassing event the young officer of-the called police-car the identity-card presented
vorgezeigt, und dann eine Blutprobe abgegeben.
and then a blood-sample given

(24) Weil der Manager den Termin vorschrieb, hat Tom Pressberger trotz heftiger Proteste des aufgebrachten
because the manger the appointment dictated has Tom Pressberger despite intense protests of-the enraged
Betriebsrats dem neuen Leiter des kleinen Vertriebsbüros den Urlaub gestrichen, und so das Arbeitsklima
workers-council the new leader of-the small sales-office the holiday canceled and thus the work-atmosphere
vergiftet.
poisoned

Experiment 2

We give one example for an experimental stimulus. As can be seem by comparison with (1), the stimuli for Experiment 2 can be derived from those of Experiment 1 by replacing the proper name with a definite NP, and turning the rest of the main clause into a relative clause that modifies this definite NP.

The example in (1) in in the condition dat = rel, adj = main. The other conditions can again be constructed by moving the dative phrase and/or the adjective phrase into the subordinate clause, as illustrated by example (7).

(1) Weil der Verkäufer den Mangel verheimlichte, hat der Kunde, der ohne langes Bedenken des
Because the seller the defect hid has the customer who without long consideration of-the
problematischen Sachverhalts dem arroganten Händler des vornehmen Autohauses den Höchstpreis geboten hat,
problematic fact the arrogant dealer of-the posh car-dealership the maximum-price offered has
einen Verlust gemacht.
a loss made