# Pitch accents create dissociable syntactic and semantic expectations during sentence processing

Constantijn L van der Burght[a,b], Angela D Friederici[a], Tomás Goucha[a,*], Gesa Hartwigsen[a,b,*]

[a] *Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany*

[b] *Lise Meitner Research Group "Cognition and Plasticity", Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany*

* Authors contributed equally

**Please cite as:**
Van der Burght, C.L., Friederici, A.D., Goucha, T., and Hartwigsen, G. (2021) Pitch accents create dissociable syntactic and semantic expectations during sentence processing. *Cognition* 212, 104702. DOI:10.1016/j.cognition.2021.104702

**Keywords**: prosody, information structure, syntax, semantics, sentence processing

**Declarations of interest:**
none

**Corresponding author:**
Constantijn van der Burght
*Department of Neuropsychology*
*Max Planck Institute for Human Cognitive and Brain Sciences*
*Stephanstraße 1A*
*04103 Leipzig, Germany*
*+49 (0)341 9940 2238*
vanderburght@cbs.mpg.de

**Abstract**

The language system uses syntactic, semantic, as well as prosodic cues to efficiently guide auditory sentence comprehension. Prosodic cues, such as pitch accents, can build expectations about upcoming sentence elements. This study investigates to what extent syntactic and semantic expectations generated by pitch accents can be dissociated and if so, which cues take precedence when contradictory information is present. We used sentences in which one out of two nominal constituents was placed in contrastive focus with a third one. All noun phrases carried overt syntactic information (case-marking of the determiner) and semantic information (typicality of the thematic role of the noun). Two experiments (a sentence comprehension and a sentence completion task) show that focus, marked by pitch accents, established expectations in both syntactic and semantic domains. However, only the syntactic expectations, when violated, were strong enough to interfere with sentence comprehension. Furthermore, when contradictory cues occurred in the same sentence, the local syntactic cue (case-marking) took precedence over the semantic cue (thematic role), and overwrote previous information cued by prosody. The findings indicate that during auditory sentence comprehension the processing system integrates different sources of information for argument role assignment, yet primarily relies on syntactic information.

## Highlights

- Two experiments on the influence of prosodic focus marking on sentence processing
- Properties of focus probed in tasks using sentence comprehension or completion
- Focus, marked by pitch accents, established syntactic and semantic expectations
- Only syntactic expectations interfered with sentence comprehension when violated
- Clear hierarchy in processing when competing cues from multiple domains are present

## 1. Introduction

Language comprehension is guided by various types of linguistic information. Previous work shows that auditory sentence processing is facilitated by expectations established by syntactic, semantic, as well as prosodic cues. One type of prosodic cue is the pitch accent, which gives prominence to a particular part of the sentence through an increase in pitch and intensity (Grabe, 1998). In written form, the sentence “John kissed Mary, not Peter” is ambiguous concerning the role of Peter: either he did not kiss Mary, or he was not kissed by John. In such cases, pitch accents can be crucial for sentence comprehension. Realising a pitch accent on either “John” or “Mary” places one of the words in so-called focus. This determines which arguments in the sentence are contrasted with each other (Rooth, 1992): either John and Peter, or Mary and Peter. Thereby the pitch accent clarifies the role “Peter” occupies in the otherwise syntactically ambiguous sentence (i.e., the pitch accent establishes *who did what to whom*). It has been suggested that the two elements that are in contrastive focus are interpreted to have parallel syntactic roles (Carlson, Dickey, Frazier, & Clifton, 2009), (Carlson, 2015). In turn, these parallels influence the interpretation of the noun phrase “Peter”, which occurs in that part of the sentence where important information is omitted, a so-called ellipsis structure (see Winkler (2019) for a review).

In sum, pitch accents, by marking contrastive focus, can draw parallels between constituents that occupy the same syntactic role. This implies that after perceiving the first focused constituent in a sentence, a certain expectation about the upcoming constituent

may be established. How different types of linguistic information interact to form these expectations is unclear. The current study investigates this interaction by exploring which expectations are formed when pitch accents highlight constituents that contain overt syntactic and semantic cues. Specifically, we asked if syntactic and semantic expectations can be dissociated, and furthermore, which type of information listeners rely on when competing cues from multiple domains are present.

There are several ways in which pitch accents can cue syntactic structure. First, they can resolve attachment ambiguities, as has been shown by several early studies on the interaction between pitch accents and syntactic structure (Schafer, Carlson, Clifton, & Frazier, 2000; Schafer, Carter, Clifton, & Frazier, 1996). For example, in "the propeller of the plane that the mechanic was so carefully examining…", a pitch accent on either "propeller" or "plane" helps to clarify what the mechanic was examining, something that is ambiguous without focus-marking (Schafer et al., 1996). It is therefore supposed that ambiguous sentence parts are likely to be attached to the sentence element that bears focus (but see Lee & Watson (2010) for an alternative explanation). Second, as discussed above, by assigning contrastive focus, pitch accents can mark parallels between constituents and influence the interpretation of their syntactic role (Carlson, 2001).

Importantly, it has been argued that the disambiguating effects of prosody are in part predictive. Several eye-tracking studies have shown that listeners anticipate a certain syntactic structure as a result of a prosodic cue (Nakamura, Arai, & Mazuka, 2012; Weber, Grice, & Crocker, 2006b). For example, Weber *et al.* (2006b) demonstrated this using sentences such as (in German) "The cat possibly hunts the dog". Because of the relatively free word order of German—meaning the object can precede the subject—this sentence contains a temporary ambiguity: "the cat" can be both subject and object of a sentence until the determiner of "the dog" is perceived. This is because "the cat" carries the syntactic gender feminine, which is not case-marked with an unambiguous form (nominative case: *die/the*; accusative case: *die/the*). The role of the noun phrase can only be disambiguated by clear case-marking of a second determiner, as in the masculine noun phrase "the dog" (nominative: *der/the*; accusative: *den/the*), causing *der/the dog* to be the subject and *den/the dog* to be the object of the sentence. However, the distribution of pitch accents on

the words “cat” and “hunts” can mark the correct interpretation of the sentence as well: a pitch accent on “cat” favours an object interpretation of *the cat*, whereas an additional pitch accent on “hunts” favours a subject interpretation. In this way, the pitch contour of sentences in which the object precedes the subject differs from those in which the subject comes first. Indeed, depending on the prosodic structure of a given sentence, listeners showed increased anticipatory eye movements to the correct interpretation in a visual scene (Weber, Grice, & Crocker, 2006b). This shows that pitch accents can influence the analysis of syntactic structure *before* additional disambiguating input has been observed (see Snedeker & Trueswell (2003) for a similar experiment using prosodic boundaries).

Aside from cueing syntax, pitch accents play an important role in the semantic domain. By marking focus, prosody forms a direct link with the information structure of a sentence. The information structure guides the listener to what is new or important in a sentence. Focus, which can be marked by pitch accents, gives prominence to sentence elements, highlighting the difference between new and given information (Jackendoff, 1972). Focus-marking is also thought to trigger semantic alternatives (reviewed in Gotzner & Spalek (2019)). For instance, in a sentence such as “Anna bought [BANANAS]”, the listener automatically considers what else Anna could have bought or did not buy (capital letters indicate focus-marking by a pitch accent). The set of alternatives that becomes activated must share semantic features with the focused constituent—although the scope and time course of this pre-activation are debated (Braun & Tagliapietra, 2010; Husband & Ferreira, 2016).

What is undisputed, however, is that this activation of semantic alternatives occurs in a predictive manner. Several eye-tracking studies have shown that after perceiving focus, participants fixate at items within a visual scene that are semantically appropriate given the focus context (Ito & Speer, 2008; Karimi, Brothers, & Ferreira, 2019; Watson, Tanenhaus, & Gunlogson, 2008; Weber, Braun, & Crocker, 2006a). For example, after having heard the instruction “Click on the purple scissors”, the follow-up instruction “Now click on the RED...” prompted listeners to look at *red scissors* rather than a different red object (Weber, Braun, & Crocker, 2006a). Here, the pitch accent marks *red* as novel information, implying that the object itself is known or given (and therefore will be scissors). Put differently, a pitch accent

on the colour adjective places a semantic restriction on the intended referent. Crucially, these fixations are "anticipatory", i.e. initiated prior to the occurrence of the target word in the auditory stimulus, pointing towards a predictive capacity of focus in the semantic domain.

Taken together, it has been shown that prosody can have a predictive function in sentence processing, both syntactically and semantically. Furthermore, in a sentence such as "JOHN kissed Mary, not PETER", the two contrastively focused arguments occupy parallel roles. This implies that after encountering the first of these constituents, there may be a certain expectation about the second, parallel one. However, it is unclear whether listeners form these expectations implicitly or explicitly, and it remains to be shown whether the contributions of syntactic and semantic information can be dissociated. Furthermore, it is unclear to what extent participants rely on syntactic and semantic cues when several contradictory indicators of sentence structure are present. To address these questions, we used focus-marking to create sentences of the type "Yesterday, the policeman arrested the thief, not the murderer" (translated from German). Realising a pitch accent on either the first (Fig 1A) or second noun (Fig 1B) resulted in the variants A and B of that sentence. Note that the noun phrases in the German sentences are marked by case (nominative [NOM] and accusative [ACC]), and that focused noun phrases are indicated with CF (contrastive focus).

A. Yesterday, [the$^{NOM}$ POLICEMAN]$_{CF}$ arrested the$^{ACC}$ thief, not [the$^{NOM}$ INSPECTOR]$_{CF}$

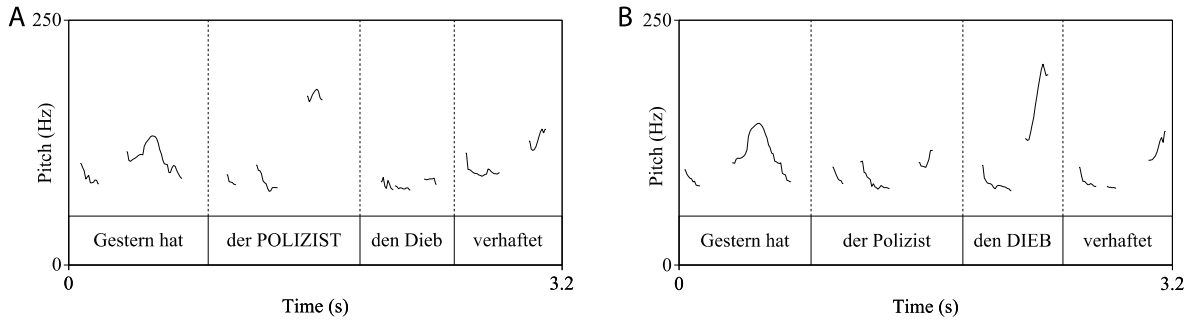
- *Gestern hat [der POLIZIST] den Dieb verhaftet, nicht [der KOMMISSAR]*

B. Yesterday, the$^{NOM}$ policeman arrested [the$^{ACC}$ THIEF]$_{CF}$, not [the$^{ACC}$ MURDERER]$_{CF}$

- *Gestern hat der Polizist [den DIEB] verhaftet, nicht [den MÖRDER]*

From a language theoretical point of view, different syntactic analyses of the ellipsis site of this particular sentence structure have been proposed. Considering ellipsis structures in general, most theories (e.g., Merchant, 2001) assume that the ellipsis contains a syntactic structure that remains unpronounced (although alternative nonstructural approaches to ellipsis have been proposed, e.g., by Ginzburg and Sag (2000) or Culicover & Jackendoff (2005); see Merchant (2018) for a recent review of this debate). The fact that in languages such as German the ellipsis structure carries case marking has been taken as evidence for

196 the existence of a resumptive structure at the ellipsis site (Ross, 1969). For the particular
197 sentence structures used in the current study, it is implied that the structure of the main
198 clause is recapitulated, but some of the constituents (the lexical verb and the noun phrase
199 that is not in focus) are redundant and therefore deleted ("Yesterday the policeman arrested
200 the thief, and the policeman did not arrest the murderer"). Prosodically, the noun phrase
201 that remains at the ellipsis site must bear a contrastive pitch accent (Winkler, 2019).
202 Importantly, these theories do not assume complexity differences between the subject and
203 object ellipsis variants (Stolterfoht, 2005), as is supported by experimental work in English
204 (Carlson, 2002). An interpretative bias between the two structures does exist, with the
205 object focus condition (*Yesterday, the policeman arrested [the THIEF]*) being the default
206 focus structure (Stolterfoht, Friederici, Alter, & Steube, 2007).
207




208
209 **Figure 1**: Pitch contours illustrating the difference between subject-focus and object-focus in
210 the example sentence *Yesterday, the policeman arrested the thief* ("Gestern hat der Polizist
211 den Dieb verhaftet"). The noun phrase that is placed in contrastive focus bears a pitch
212 accent (indicated by capital letters), whereas it is deaccented in the other condition.
213

214 To dissociate syntactic and semantic processes, we included explicit cues in both
215 domains. As syntactic cue, we made use of the German case system, since German speakers
216 have been consistently shown to follow the syntactic cues provided by the case marking of
217 the determiners (E. Bates & MacWhinney, 1989; Bornkessel-Schlesewsky et al., 2011). As
218 described before, sentence elements in German are mostly free to occupy different
219 positions along the sentence. However, the overt case marking of both determiners and
220 pronouns determines the syntactic function of nominal constituents, thus establishing
221 sentence structure. In our experiments, we used the masculine determiners for the critical
222 syntactic conditions because of their unambiguous case marking: each of the four cases has

a specific masculine singular form different from the other three forms. In particular, we used the nominative subject form “der” and the accusative object form “den”. Various paradigms have used the contrast of these two forms of the determiner to investigate syntactic processing (Clahsen & Featherston, 1999; Kamide, Scheepers, & Altmann, 2003; Strotseva-Feinschmidt, Cunitz, Friederici, & Gunter, 2015), and there is also evidence that, during development, German speakers start to mainly rely on case to identify the subject and object roles in the sentence (Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008). Therefore, we used grammatical case of the determiners in our experimental items as a syntactic marker of the subject [NOM] vs object role [ACC]. As semantic cue, we made use of thematic role typicality: the notion that a verb is associated with a set of thematic roles, corresponding to the participants in an event (Jackendoff, 1972). For example, the verb “to arrest” has typical agents (doers, e.g. a policeman) and patients (undergoers, e.g. a thief) that participate in the action. Using this combination of cues, we tested the hypotheses that a pitch accent on either the subject or object noun phrase of the main clause should establish expectations concerning the syntactic and semantic content of the ellipsis structure.

We examined the characteristics of these expectations in two experiments. In Experiment 1, we tested if these expectations can be probed *implicitly*. We should then be able to find evidence of inhibited processing in case these expectations are violated. To tease apart the syntactic and semantic components of these expectations, we manipulated the syntactic and semantic (mis)match between the upcoming constituent in the ellipsis part and the expectations formed in the main clause. In this experiment, listeners were then asked whether they interpreted the different noun phrases of the sentence as subject or object. If mismatching cues between two focused constituents resulted in delayed responses, we can argue that the noun phrase in the ellipsis part violated a syntactic or semantic expectation established by the pitch accent. Experiment 2 investigated if listeners form an *explicit* expectation, in which case we should find evidence of prosodic focus-marking when directly probing the listener’s preferred continuation of a sentence. To test whether participants formulated an explicit prediction, participants in this experiment completed an auditory sentence, which was cut before the second focused constituent was produced (“Yesterday, the POLICEMAN arrested the thief, not…”), by selecting the case of

255 the determiner and the role of noun. Together, these two experiments enabled us to
256 investigate the expectations that pitch accents establish, and to what extent they can
257 manipulate the interpretation of an ellipsis structure. By highlighting constituents that
258 contain a syntactic as well as semantic cue, we could assess if syntactic and semantic
259 processes can be dissociated within these expectations. Finally, considering that syntactic
260 cues and thematic roles interact (Trueswell, Tanenhaus, & Garnsey, 1994), we asked to what
261 extent listeners rely on syntactic and semantic components when multiple indicators of
262 sentence structure are available.
263

## 2. Experiment 1

### *2.1. Methods*

The design and analysis plan of this experiment were preregistered at the Open Science Framework (Experiment 1: https://osf.io/94bp5). Experiment 1 involved a sentence comprehension task with a 3x2x2 factorial within-subject design with the factors focus mismatch type (baseline; semantics; syntax), focus position (subject; object) and target of comprehension question (main clause; ellipsis). Raw data and analysis scripts can be found at https://osf.io/v5xga/.

#### *2.1.1. Participants*

36 healthy native German speakers (20 female; age *M* = 23.8 years, *SD* = 4.0, range 18-34) were included in the analysis. Participants (Oldfield, 1971) were recruited from the database of the Max Planck Institute for Human Cognitive and Brain sciences. All participants had normal or corrected-to normal vision. We chose to invite only right-handed participants in light of a planned follow-up study involving non-invasive neurostimulation, for which we required a uniform sample of right-handers. Exclusion criteria were hearing loss or professional musical training. One participant was excluded from the analysis because of incorrect handedness information. The experiment was approved by the ethics committee of the University of Leipzig, and all participants gave written consent prior to participation.

To determine our sample size, we ran a power analysis using the powerSim function of the simR package in R on data from an independent sample tested in a pilot version of the experiment (n=7). We tested for the interaction term *focus mismatch type* x *comprehension question target* from our original hypothesis, running 25 simulations in 36 participants. We determined this initial sample size of 36 to have a minimum of 1600 observations per cell (Brysbaert & Stevens, 2018). Our stimulus set consisted of 48 items, leading to an estimated sample size required of at least 34 (1600/48=33,33), to which we added 2 to achieve a full balancing-out of our design. These simulations yielded an estimated power of above 99%. This suggests that a smaller sample size would achieve sufficient power, however, to avoid going below the minimum number of observations recommendation by Brysbaert & Stevens (2018), we determined our required sample size at 36.

### *2.1.2. Stimulus design*

In our stimulus sentences (in German), one out of two constituents in a first clause was placed in contrastive focus with a third constituent in a second, elliptical clause (as exemplified previously in sentences A and B; analogous to Stolterfoht, Friederici, Alter, & Steube (2007). A pitch accent (indicated with capital letters in the examples below) marked whether focus was on the subject (1) or the object noun phrase (2). To tease apart the syntactic and semantic components of the expectations created by focus, the noun phrases contained specific syntactic information (case marking of the determiner) and semantic information (thematic role of the noun). In the ellipsis structure that followed, a third noun phrase occurred that corresponded grammatically and thematically to the focused noun phrase in the main clause (baseline condition).

(1) Yesterday, [the$^{NOM}$ POLICEMAN]$_{CF}$ arrested the$^{ACC}$ thief, not [the$^{NOM}$ INSPECTOR]$_{CF}$

- *Gestern hat [der POLIZIST] den Dieb verhaftet, nicht [der KOMMISSAR]*

(2) Yesterday, the$^{NOM}$ policeman arrested [the$^{ACC}$ THIEF]$_{CF}$, not [the$^{ACC}$ MURDERER]$_{CF}$

- *Gestern hat der Polizist [den DIEB] verhaftet, nicht [den MÖRDER]*

In (1) the determiners of the two contrasted noun phrases are in nominative case, and both nouns are typical agents of the verb “to arrest”. In (2) the contrastive constituents are case-marked accusative and typical patient nouns.

To form syntactic and semantic mismatches between the two focused constituents, we created combinations with mismatching grammatical case and thematic roles. In the condition with a syntax-focus mismatch (3 and 4), the grammatical case of the determiner in the ellipsis structure mismatches the focused constituent in the main clause (nominative vs. accusative).

(3) Yesterday, [the$^{NOM}$ POLICEMAN]$_{CF}$ arrested the$^{ACC}$ thief, not [the$^{ACC}$ INSPECTOR]$_{CF}$

- *Gestern hat [der POLIZIST] den Dieb verhaftet, nicht [den KOMMISSAR]*

(4) Yesterday, the$^{NOM}$ policeman arrested [the$^{ACC}$ THIEF]$_{CF}$, not [the$^{NOM}$ MURDERER]$_{CF}$

- *Gestern hat der Polizist [den DIEB] verhaftet, nicht [der MÖRDER]*

In the condition with a semantics-focus mismatch (5 and 6), the thematic role in the ellipsis structure mismatches the focused noun in the main clause (typical agent vs. patient).

(5) Yesterday, [the$^{NOM}$ POLICEMAN]$_{CF}$ arrested the$^{ACC}$ thief, not [the$^{NOM}$ MURDERER]$_{CF}$

- *Gestern hat [der POLIZIST] den Dieb verhaftet, nicht [der MÖRDER]*

(6) Yesterday, the$^{NOM}$ policeman arrested [the$^{ACC}$ THIEF]$_{CF}$, not [the$^{ACC}$ INSPECTOR]$_{CF}$

- *Gestern hat der Polizist [den DIEB] verhaftet, nicht [den KOMMISSAR]*

The experimental items consisted of verb-argument combinations with clear agent-patient relationships. All nouns were required to be masculine to enable the overt morphosyntactic marking of grammatical case of the determiners (in German, the nominative and accusative case of feminine and neuter determiners share surface form). Furthermore, we excluded nouns with different forms for the nominative and accusative case, expressing a morphosyntactic ending in the accusative form (for example, the word *student* in German is "Student" in the nominative, but "Studenten" in the accusative case). This inflection is one of the rare expressions of nominal case in German, since case in German is mainly expressed at the determiners and adjectives. By excluding such forms, we ensured that in our experiment case was marked solely by the determiner. The number of syllables of the nouns that belonged to the same verb was matched as closely as possible.

To investigate whether participants have an intrinsic bias of the sentences toward object or subject contrast, we carried out an online normative study rating the two baseline conditions. The experiment was programmed in Psychopy, version 2020.1.2 (Peirce et al., 2019) and run online via Pavlovia. Participants (N=40) listened to all baseline sentences (48 stimuli per condition) and were subsequently asked to rate how much they liked the sentence ("Wie gefällt dir der Satz?") on a scale from 1 ("gar nicht"/not at all) to 7 ("sehr"/very much) using their keyboard. We choose to investigate a possible structural bias between the two focus constructions by probing a general evaluation of the sentences, to avoid that participants were drawn to either syntactic, semantic, or acoustic stimulus properties in making their judgement. Results were analysed by running a cumulative link mixed model using the clmm function of the package ordinal (version 2019.12-10) in R (version 4.0.2; R Core Team (2020)). We modelled the rating scores in function of focus as

fixed effect. The random effects structure existed of subject-wise random intercepts and slopes for the factor focus (Barr, Levy, Scheepers, & Tily, 2013). A likelihood ratio test comparing the full model to the null model (omitting the factor Focus) did not demonstrate a significant effect of Focus ($LR$ = 0.80, $p$ > .05). A histogram with response distributions for the subject and object focus stimuli is presented in Supplementary Figure 1, with detailed model output provided in Supplementary Table 1. The experiment and relevant code are available upon request at https://gitlab.pavlovia.org/vanderburght/norming.

The semantic properties of the materials were evaluated in a normative study on a separate sample (n=40) based on Ferreira (2003). To assess the semantic-thematic relationships between the verbs and their noun phrase arguments, all verbs were presented with an agent and patient in plausible and implausible order. The items were divided over four lists, such that each participant rated each verb twice: with one agent-patient pair in a plausible sentence (e.g. *The policeman arrested the murderer*) and a different pair in an implausible sentence (e.g. *The thief arrested the detective*). Participants were instructed to carefully read the sentences and rate them on a scale from 1 ("extremely implausible") to 6 ("extremely plausible"), with examples provided. From an initial set of 73 items the 48 items with the largest plausible and implausible difference were selected. These 48 verb-argument combinations had a clear thematic role assignment, with the implausible versions rated less plausible than their plausible counterparts (plausible: $M$ = 5.33, $SD$ = 0.40; implausible: $M$ = 1.57, $SD$ = 0.55).

### *2.1.3. Stimulus construction*

A professional native German speaker was recorded producing two variants of 48 critical items (listed in Supplementary Table 5). The speaker was instructed to realise a pitch accent on either the subject (Fig. 1A) or object (Fig. 1B) of the main clause. More specifically, the speaker was instructed to realise a low tonal target followed by a steep rise to the pitch maximum (L+H*), since in German contrastively focused constituents are typically marked by this type of pitch accent (Weber et al. 2006a; Braun & Tagliapietra, 2010). At the sentence-final position, a filler noun phrase was produced that was later removed. The sentence-final noun phrases were taken from separate recordings: a typical agent (a and c) or typical patient (b and d) of the verb (in this case "to arrest"), combined with a determiner in either

nominative (a and b) or accusative case (c and d). These sentence-final nouns all carried a contrastive pitch accent.

a) … the$^{NOM}$ INSPECTOR

b) … the$^{NOM}$ MURDERER

c) … the$^{ACC}$ INSPECTOR

d) … the$^{ACC}$ MURDERER

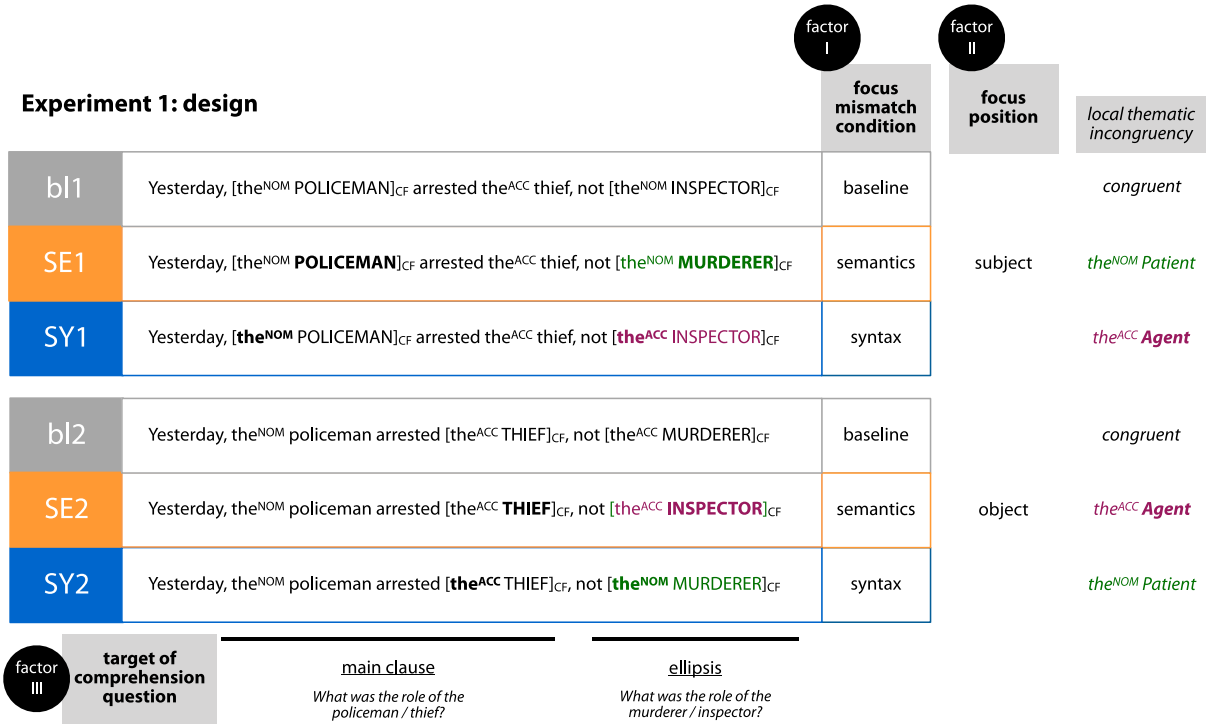



**Figure 2:** Design overview of Experiment 1. The two factors focus mismatch condition and focus position resulted in six sentences. Each sentence could be probed by a comprehension question related to the noun phrase at the main clause or ellipsis (experimental factor 3). Violating sentence elements in bold typeface. Sentence-final determiner-noun pairs are colour-coded separately (see Results). Pitch accents are indicated by capital letters. The outer right column displays the local thematic incongruencies present at the ellipsis site, which were necessary to create the focus mismatch conditions. bl = baseline; se = semantic; sy = syntactic; CF = contrastive focus.

The items in a)-d) enabled us to create combinations in which the two focused constituents either had corresponding grammatical case (determiners) and thematic role typicality (nouns) or carried mismatching syntactic or semantic information. The cross-splicing procedure ensured that the comparisons between conditions of interest involved materials

that were acoustically identical, and the speaker never had to produce sentences containing mismatching syntactic or semantic information. Participant debriefings during the pilot stage of the experiment ensured that the audio manipulation was not audible and that all stimuli were perceived as natural. A sound wave and spectrogram of an example stimulus can be found in Supplementary Figure 2. Figure 2 provides an overview of the experimental conditions.

Recordings were made in a sound-attenuating chamber (IAC – I200 series, Winchester, United Kingdom) and the digitised speech signals (sampling rate 44.1kHz; resolution 16 bits) were adjusted to the same root mean square amplitude using Praat (Boersma & Weenink, n.d.). In the same programme, sound files were manually cut and subsequently concatenated using a custom-made script.

#### 2.1.4. *Procedure*

Participants performed a sentence comprehension task (Fig 2 & 3). At trial onset, a white fixation cross was presented which turned red 200 ms prior to auditory onset to alert the participant. The auditory stimulus was followed by the comprehension question and two visually presented answer options. Participants responded via button press with the right index or middle finger. Subsequently, a fixation cross was presented for approximately 2 s until the next trial started.

A comprehension question probed how listeners interpreted the sentence. The comprehension question could target one of the two noun phrases in contrastive focus: these questions—probing either the focused noun phrase in the main clause or the focused noun phrase in the ellipsis—occurred equally often and were presented counter-balanced across conditions. We chose to probe both focused constituents across trials to ensure that listeners would be equally attentive to both the main clause and ellipsis part of the sentences. Listeners were asked what role a certain participant played in the action described in the sentence*: "What was the role of the policeman?"* (in subject-focus trials) or *"What was the role of the thief?"* (in object-focus trials). They indicated whether the policeman/thief was doer or undergoer of the action *("has arrested"* or *"was arrested"*). If the noun phrase in the ellipsis structure was probed (*"What was the role of the inspector/murderer?"*) the response options were *"has not arrested"* or *"was not arrested".*

The assignment of the active/passive answer options to the response buttons was counter-balanced between subjects.

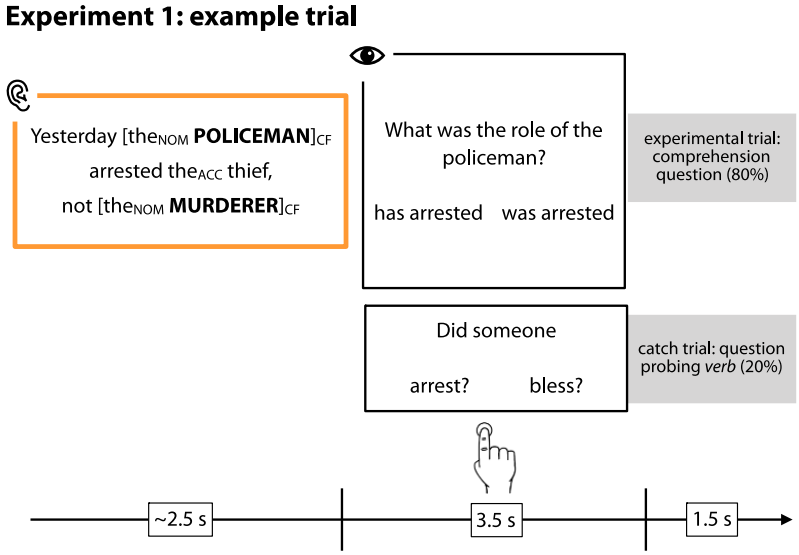



**Figure 3:** Example trial of Experiment 1 – sentence comprehension paradigm. Experimental trials contained comprehension questions probing one of the two contrastively focused noun phrases. Catch trials probed the verb.

The trial sequences were pseudo-randomised with the following constraints: each item (verb) was presented once in each block of 48 trials; the same focus mismatch conditions, focus position, and the target of the comprehension questions (probing either main clause or ellipsis part of the sentence) were not repeated more than twice. To draw the participants' attention to the semantic-thematic content of the verb-argument structure rather than merely to the three noun phrases, catch trials were included (amounting to 20% of the total number of trials) which probed the verb of the auditory stimulus (e.g. *Did someone… arrest / instruct?*). This resulted in the following composition of the stimulus set: of all items, 46.67% were congruent (26.67% experimental items + 20% filler items used in the catch trials), 26.67% contained mismatching syntactic information, and 26.67% contained mismatching semantic information. The experiment lasted for approximately 52 minutes including 5 breaks, the duration of which was self-timed. A short practice session preceded the experiment, mirroring the main experiment but consisting of different stimuli.

Participants sat in a sound-attenuated chamber and listened to the auditory stimuli over headphones. Visual stimuli were presented on a screen (Sony Trinitron Multiscan 300GS, Sony Corporation) and responses were given on a response-box placed on their lap.

Stimulus presentation and response collection was controlled via Presentation (Neurobehavioural Systems, Inc., Albany, CA, USA).

#### 2.1.5. *Data analysis*

Response times were analysed using a Linear Mixed Model (Baayen, Davidson, & Bates, 2008). Upon visual inspection, response times were log-transformed to approach a normal distribution. The proportion of responses *active/passive* were analysed using a logistic Generalized Linear Mixed Model (Baayen et al., 2008). In both models, we included the factors *focus mismatch condition, focus position, comprehension question target*, and their interaction as fixed effects. The three-level focus mismatch condition factor was dummy coded with the semantic condition being the reference category; the two-level factors were sum-coded. Contrary to our a-prior hypothesis that only the factors *focus mismatch condition* and *comprehension question target* would interact, visual inspection of the response times (see Fig 4 & 5) motivated us to consider a three-way interaction as the most appropriate way to model the data. We aimed to include a maximal random effects structure (Barr et al., 2013). However, due to convergence issues, we simplified the random effects structure until the models converged, by removing the interaction terms and finally the main effects, first for item and then for participant (for the main effects, we prioritised inclusion of the factor *focus mismatch condition*). This led to the use of an intercept-only model in the Linear Mixed Model (reaction times) and inclusion of by-participant random slopes for the factor *focus mismatch condition* in the Generalized Linear Mixed Model (for the proportion of responses).

We tested the effect of the three-way interaction using a likelihood ratio test comparing the full model with a reduced model lacking the interaction term (Barr et al., 2013); Dobson & Barnett, 2008). Pair-wise follow-up comparisons were done by calculating estimated marginal means (Searle, Speed, & Milliken, 1980) using the package emmeans (Lenth, Singmann, Love, Buerkner, & Herve, n.d.). The models were fitted in R (version 3.6.0; R Core Team (2019)) using the functions lmer and glmer of the package lme4 (version 1.1-21; Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). We used raincloud plots (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2018) for visualisation of the response times, to show both summary statistics and the response distributions per condition.

## 2.2. **Results**

In the response times, we found a significant interaction between focus mismatch condition, focus position, and comprehension question target ($\chi^2(2)=30.63$, $p<.0001$). The same three-way interaction was significant in the analysis of the response proportions ($\chi^2(2)=29.71$, $p<.0001$). Response behaviour to comprehension questions targeting the main clause of the stimuli is shown in Figure 4. Behavioural results when targeting the ellipsis with the comprehension question are shown in Figure 5. The estimated fixed and random effects are shown in Tables A1 (reaction times) and A2 (response proportions).

### *2.2.1. Comprehension question probing the main clause*

In the interpretations of the main clause, planned pair-wise comparisons showed significant increase in response times of the syntactic condition as compared to the semantic and baseline conditions (Fig 4A and Supplementary table 2). This was the case after subject focus – syntactic vs semantic: $t(10225)=-3.820$, $p=.001$; syntactic vs baseline: $t(10225)=-4.847$, $p<.001$) – and after object focus – syntactic vs semantic: $t(10225)=-2.689$, $p=.036$; syntactic vs baseline: $t(10225)=-3.318$, $p=.006$ – (note that the high number of degrees of freedom is due to single-trial information on which the estimated marginal means are based).

Importantly, the sentence material that participants were asked to interpret in the main clause was identical in all conditions: *the*$^{NOM}$ *policeman* in case of a subject-focus stimulus, and *the*$^{ACC}$ *murderer* after an object-focus stimulus. The sole difference between the conditions was the type of mismatch (semantic or syntactic) that followed in the ellipsis part of the sentence. These violations are reflected in the response times, with the syntactic mismatch leading to an additional processing cost. The proportions of subject/object judgements (Fig 4B) were not affected by these violations: analysing the proportion of responses that correctly interpreted the syntactic and semantic cues of the main clause, there were no significant differences in the pair-wise comparisons between the response proportions of each condition (see Supplementary Table 2: Response proportions). Finally, to explore the development of these effects along the experiment, we provide descriptive statistics of the behavioural measures across time bins in Figure A1. These suggest a reduction in the effect of the mismatch between focus and syntax as compared to baseline over the course of the experiment.

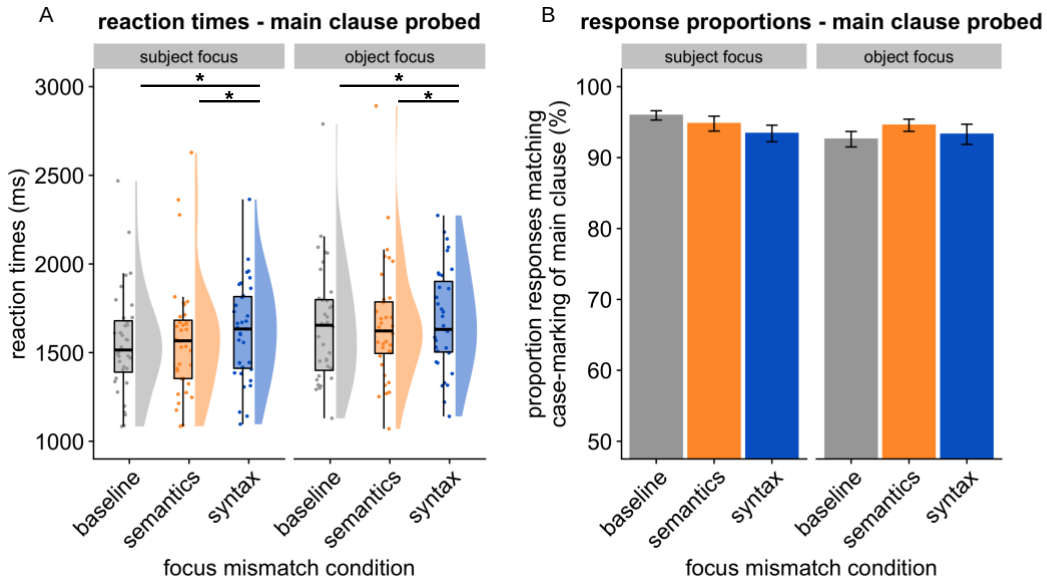



**Figure 4**: Reaction time (A) and response proportions (B) for the comprehension questions that probed the main clause (Experiment 1). Error bars indicate ±1 SEM. Asterisks mark planned pair-wise comparisons with *p*-values smaller than .05 (Bonferroni-Holm-corrected).

### *2.2.2. Comprehension question probing the ellipsis*

In the responses at the ellipsis site (Fig 5), we did not find the same pattern of results as was found in the main clause responses. That is, we found no evidence of syntactic expectations that had been generated by the prosodic cue in the main clause (Fig 4A). Rather, the response times differences of the semantics and syntax conditions depended on whether focus in the main clause was on the subject or object noun phrase: responses were faster in the syntax condition as compared to the semantics condition after subject focus ($t(10225)=5.564$, $p<.001$), whereas after object focus responses were slower in syntax than in semantics ($t(10225)=-4.103$, $p<.001$). Pairwise comparisons are presented in Supplementary table 2.

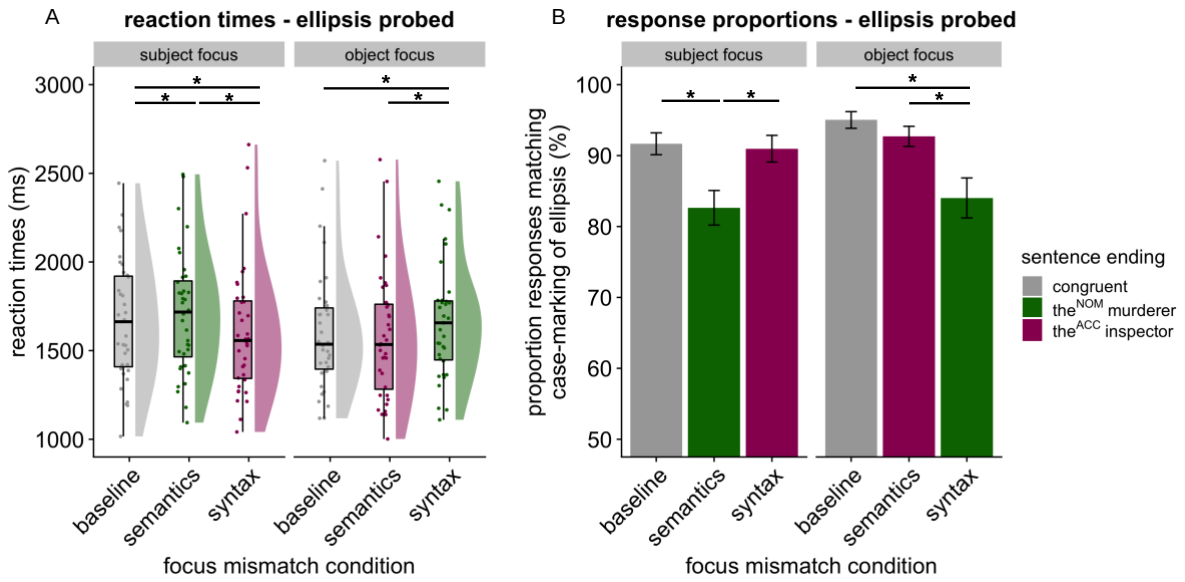



**Figure 5**: Reaction time (A) and response proportions (B) for the comprehension questions that probed the ellipsis clause (Experiment 1). Error bars indicate ±1 SEM. Asterisks mark planned pair-wise comparisons with *p*-values smaller than .05 (Bonferroni-Holm-corrected).

Here, we need to take into account that, at the ellipsis site, participants were asked to make a judgement on the role of a noun phrase, which in itself held conflicting semantic and syntactic information except in the baseline condition: in *the$^{NOM}$ MURDERER,* a typical patient of *to arrest* was preceded by a determiner in the nominative case (cueing a subject role); in contrast, *the$^{ACC}$ INSPECTOR* is a typical agent preceded by a determiner in the accusative (cueing an object role). When considering the congruency of grammatical case and role typicality at the ellipsis, the pattern of response times shows a striking correspondence: response times were shorter when a role judgement was required on *the$^{ACC}$ INSPECTOR,* but longer when judging *the$^{NOM}$ MURDERER*. In sum, the pattern of reaction times at the ellipsis does not reflect the type of mismatch present between the two focused constituents across the sentence, but rather the local grammatical-thematic congruency of the determiner-noun pairs at the ellipsis site itself.

This interpretation is supported by the analysis of the response proportions at the ellipsis part of the sentence (Fig 5B). Firstly, participants responded according to the grammatical case of the determiner presented at the ellipsis site: *the$^{ACC}$ INSPECTOR* was interpreted as object of the sentence and *the$^{NOM}$ MURDERER* as subject, despite the conflicting semantic information. However, in the case of *the$^{NOM}$ MURDERER*, we observed

fewer responses corresponding to the case-marking cue of the ellipsis as compared to the other sentence endings: specifically, there was a significant decrease in the number of "subject" responses (a "subject" response is in line with the nominative case of the determiner). This pattern driven by sentence endings was present both after subject focus (semantics vs baseline: $z$=5.065, $p$=<.001; semantic vs syntactic: $z$=-5.005, $p$<.001) and after object focus (syntactic vs baseline: $z$=6.643, $p$<.001; syntactic vs semantic: $z$=5.524, $p$<.001).

Importantly, the role judgements made at the ellipsis site corresponded to the syntactic cue presented at the ellipsis site, regardless of whether conflicting syntactic or semantic information was focused in the main clause. This implies that, even though pitch accents can establish an expectation concerning upcoming syntactic information (as can be seen in the response times of the main clause), it is the incoming local syntactic cue that is decisive for the role judgement at the ellipsis site.

Finally, to explore the development of these effects along the experiment, we provide descriptive statistics of the behavioural measures across time bins in Figure A2. These suggest a reduction in the effects of the local syntactic-semantic incongruency in *the$^{NOM}$ MURDERER* as compared to baseline over the course of the experiment.

## 3. Experiment 2

From Experiment 1, it remained unclear whether prosodically-marked semantic information establishes expectations about upcoming sentence constituents since pair-wise comparisons between semantics and baseline were not significantly different. We therefore conducted a follow-up experiment, in which the stimuli from Experiment 1 had the final constituent removed and in which participants had to explicitly continue the sentence in a forced-choice task (see Figure 6). The removal of the sentence final constituent resulted in (I) and (II).

(I) Yesterday, [the$^{NOM}$ POLICEMAN]$_{CF}$ arrested the$^{ACC}$ thief, not …

- *Gestern hat [der POLIZIST] den Dieb verhaftet, nicht …*

(II) Yesterday, the$^{NOM}$ policeman arrested [the$^{ACC}$ THIEF]$_{CF}$, not …

- *Gestern hat der Polizist [den DIEB] verhaftet, nicht …*

Participants were then asked to listen to the beginning of the sentence and to complete the sentence. Crucially, the appropriate determiner and noun of the missing noun phrase had to

be chosen sequentially: participants first selected a case-marked determiner (syntactic completion) and then a noun (semantic completion). We created separate syntactic and semantic experimental conditions as follows, to prevent the syntactic decision from influencing the subsequent semantic decision.

In the syntactic condition, participants had to choose between two determiners marked in nominative or accusative case (*der/the*$^{NOM}$ or *den/the*$^{ACC}$). By presenting the decision on the determiner first, participants made a purely syntactic decision, without possible semantic influence from a co-occurring noun. In the semantic condition, the agent and patients were presented in their feminine versions. In German, nominative and accusative case marking of the feminine determiner *die/the* is ambiguous (representing both cases). In this way, the decision on the determiner on sentences with feminine noun phrases was meaningless. Consequently, the subsequent decision on the noun (*police officer*$^{FEM}$ or *thief*$^{FEM}$*)* was a purely semantic one, without possible influence from a preceding syntactic judgement.

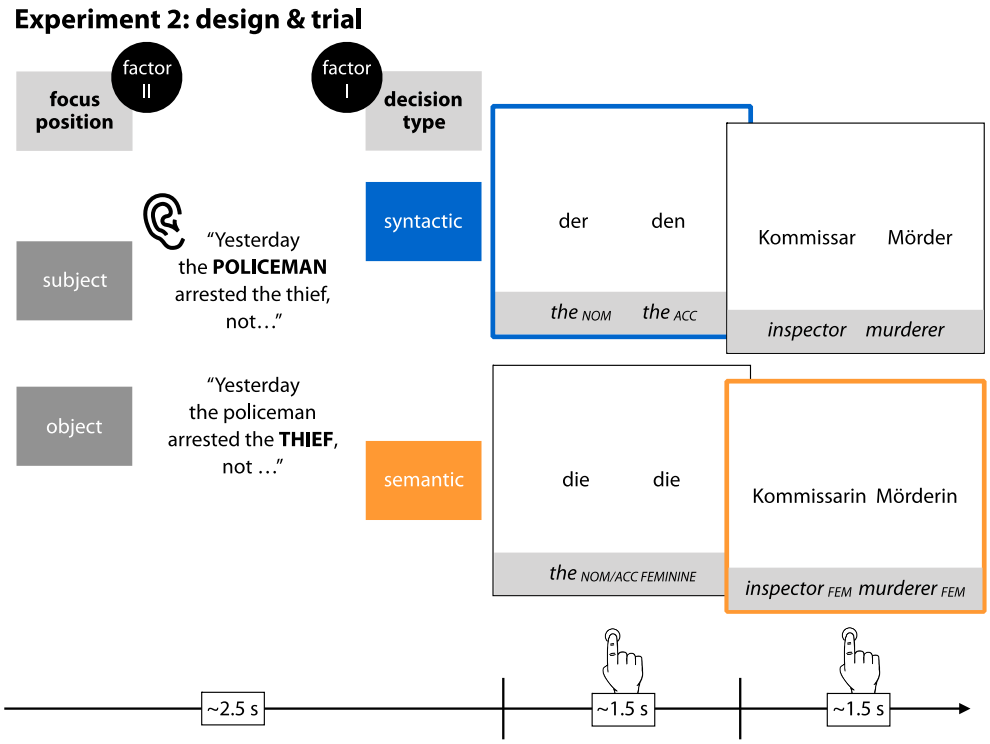



**Figure 6**: Experimental design of Experiment 2 – completion paradigm. Pitch accents are indicated by capital letters. NOM = nominative; ACC = accusative; FEM = feminine.

### 3.1. Methods

Experiment 2 involved a sentence completion task, using a 2x2 factorial within-subject design with the factors *decision type* (syntactic; semantic) and *focus position* (subject; object). Raw data and analysis scripts can be found at https://osf.io/v5xga/.

#### *3.1.1. Participants*

36 native German speakers (19 female; age *M* = 24.6 years, *SD* = 4.9, range 18-35) who had not taken part in Experiment 1 were included in the analysis. The inclusion and exclusion criteria were the same as those for Experiment 1. Eight additional data sets had to be excluded (incorrect handedness information, n=1; native language other than German, n=1; incorrect button-response pairing, n=6). We determined our sample size at 36 to remain analogous to Experiment 1, despite the difference in complexity of the design. The experiment was approved by the ethics committee of the University of Leipzig, and all participants gave written consent prior to participation.

#### *3.1.2. Stimulus description*

The auditory stimuli were the same sentence beginnings as used in Experiment 1 (resulting in sentences (I) and (II)). In this sentence completion task, participants were asked to make a syntactic judgement by choosing an appropriate continuation of the sentence (a determiner in either the nominative or accusative case, presented visually). The two nouns that were presented subsequently (a pair of a typical agent and patient of the verb in the preceding spoken stimulus) were taken from the sentence endings of Experiment 1 (the nouns from (a)-(d), see Supplementary table 5). We used feminine versions of these nouns for the semantic condition (see below).

#### *3.1.3. Procedure*

In this experiment participants performed a sentence completion task: the stimuli from Experiment 1 were cut before the noun phrase in the ellipsis part (sentences (I) and (II), modified from sentences (1) and (2)) and participants completed them by button-press in a two-alternative forced choice task. As described in the introduction to Experiment 2, participants made two consecutive decisions: they first selected a determiner and then a noun. In the syntactic condition, participants chose between two determiners marked in

nominative or accusative case (*der/the*$^{NOM}$ or *den/the*$^{ACC}$). By presenting the determiners first rather than simultaneously with the noun, participants made a purely syntactic decision, void of a possible semantic influence (see Figure 6). In the semantic condition, the agent and patients were presented in their feminine versions. As nominative and accusative case-marking of the determiner *the* is ambiguous in German, the decision on the determiner was meaningless. The subsequent decision on the noun ("*policewoman*" or "*thief*$^{FEM}$*")* was therefore a purely semantic one, without possible influence from the preceding syntactic judgement. Participants were instructed to select the determiner and noun that would complete the sentence in the way they deemed most sensible. Participants were not explicitly made aware of the meaningless choice between the feminine determiners prior to the practice trials of the experiment. However, after the practice phase, it was explained that during these trials they could respond with whichever button. These responses were not part of any further analysis. They were asked to give their response as quickly and accurately as possible.

Trial sequences were pseudo-randomised with the following constraints: stimuli with the same focus position were not repeated more than twice, and syntactic and semantic response conditions not more than three times. The assignment of the nominative/accusative and agent/patient answer options to the response buttons was counter-balanced within subjects. As in Experiment 1, each trial started with a white fixation cross which turned red 200 ms prior to the onset of the auditory stimulus. After the interrupted sentence, the two determiner options were presented visually. The two nouns were presented as soon as the response to the determiner was made (or after 1500 ms in case of a missing response). The experiment lasted for approximately 25 minutes, including 3 self-timed breaks.

#### *3.1.4. Data analysis*

Reaction times and response proportions were analysed in the same way as for Experiment 1. As fixed effects, the factors *decision type* and *focus position* and their interaction were included; both factors were sum-coded. We aimed to include a maximal random effects structure (Barr et al., 2013), however, due to convergence issues, we simplified the random effects structure until the models converged (see Experiment 1). This led to the inclusion of by-participant random slopes for the factors *decision type* and *focus position* and by-item

687 random intercepts in the Linear Mixed Model (reaction times), and an intercept-only
688 structure in the Generalized Linear Mixed Model (response proportions). To investigate
689 possible interaction effects, likelihood-ratio tests were performed comparing the full model
690 to the reduced model lacking the interaction term (Barr et al., 2013; Singmann & Kellen,
691 2019). To confirm that the pitch accent manipulation was perceived and determined
692 response patterns in the syntactic and semantic decision types, it was required that
693 participants performed above chance in all conditions. For this we used the intercept
694 estimate in the binomial model: an intercept deviating from 0 indicates that the proportion
695 of subject/object responses is not divided equally over the reference levels of the factors
696 (suggesting a deviation from chance performance). We re-leveled our fixed effects *decision*
697 *type* and *focus position* to obtain the intercepts for all four cells (subject focus, syntactic
698 decision; subject focus, semantic decision; object focus, syntactic decision; object focus,
699 semantic decision).
700 Furthermore, we performed an exploratory follow-up analysis investigating the
701 inherent bias of individual participants to choose nominative/accusative determiners or
702 agent/patient-like nouns. We employed methods from signal detection theory (Macmillan &
703 Creelman, 1991) to dissociate sensitivity to the prosodic manipulation (d-prime) and
704 response bias. To this end, we treated the subject-focus trials as 'signal' and object-focus
705 trials as 'noise'. Responses congruent with subject and object roles were coded as 'hits' and
706 'correct rejections', respectively. Incongruent responses were coded as 'misses' (subject
707 focus) and 'false alarms' (object focus) (see Meyer et al. (2016) for a similar approach).
708
709 **3.2. Results**
710 We found a significant interaction between the factors decision type and focus position in
711 the response times ($\chi^2(1)=21.19$, $p<.001$) as well as the response proportions ($\chi^2(1)=40.08$,
712 $p<.001$). More importantly, participants performed above chance in all conditions, indicating
713 that their syntactic and semantic judgements depended on the focused constituent in the
714 main clause: after subject focus (sentence I), participants preferred to continue the sentence
715 with a determiner in the nominative case ($M=71.0\%$, $SE=4.1\%$, $z=5.77$, $p<.001$) and an agent-
716 like noun ($M=76.2\%$, $SE=3.2\%$, $z=7.25$, $p<.001$). After object focus (sentence II), we saw the
717 opposite pattern: accusative-marked determiners were preferred ($M=76.0\%$, $SE=3.5\%$,
718 $z=7.20$, $p<.001$) as well as patient-like nouns ($M=68.7\%$, $SE=3.7\%$, $z=5.13$, $p<.001$). This shows

that focus established an expectation about syntactic structure as well as semantic content of the upcoming clause (Fig. 7B). The estimated fixed and random effects of this experiment are shown in Tables A3 (reaction times) and A4 (response proportions).

Experiment 2 shows that participants formed a syntactic expectation that could be probed *explicitly*, since their preferred sentence continuation was syntactically congruent with the focused constituent they had perceived. This evidence goes in line with our result of Experiment 1, in which a mismatch between syntactic information in the main clause and in the ellipsis led to an inhibited interpretation of the role of the focused noun phrase in the main clause. In other words, results from both experiments suggest that the focused constituent establishes an expectation concerning the syntactic structure of the ellipsis: this is suggested by delayed responses in case this expectation is violated (Experiment 1) and by the preference for determiners that are syntactically congruent with the focused constituent of the main clause (Experiment 2). In addition, Experiment 2 shows that focus can indeed establish an expectation about the semantic content of an upcoming clause, at least when explicitly probed: participants based their agent/patient preference on whether they had perceived a focused subject or object in the main clause. Specifically, there was a preference for typical agent nouns after subject focus sentences and typical patients after object focus. Since these semantic predictions did not cause an increase in response times in the main clause of Experiment 1, this indicates that although pitch accents can establish semantic expectations, they are not sufficiently strong to lead to additional processing cost in case they are violated.

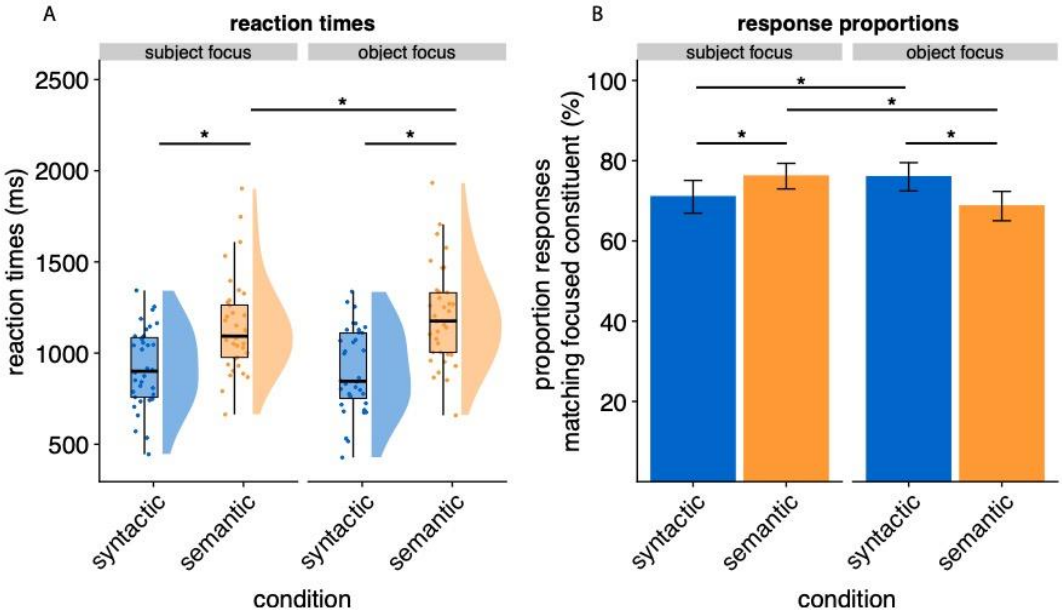



**Figure 7:** Reaction time (A) and response proportions (B) for Experiment 2. Error bars indicate ±1 SEM. Asterisks mark planned pair-wise comparisons with *p*-values smaller than .05 (corrected using the Bonferroni-Holm method).

The decreased proportion of focus-congruent responses in syntactic decisions after subject-focus as compared to object-focus ($z$=-3.607, $p$<.001) and in semantic decisions after object-focus as compared to subject-focus ($z$=5.351, $p$<.001) may reflect that participants had an overall preference for accusative determiners and agent-like nouns, respectively (see Supplementary Table 4 for all planned pair-wise comparisons). A similar pattern was found in the response times (Fig 7A), where semantic decisions where faster after subject as compared to object focus, indicating a preference for agent-like nouns ($t$(108.546)=-4.329, $p$<.001). In the syntactic decisions, response times suggested the opposite pattern, although the difference between subject vs. object focus was not significant ($t$(107.978)=1.561, $p$<.121). The possibility of opposite subject vs. object preferences in the syntactic and semantic domains, in combination with the between-subject variability in the response patterns, led us to conduct an exploratory analysis using signal detection theory methods (Figure 8). This analysis enabled us to distinguish between sensitivity to the prosodic manipulation and a possible response bias. From the plots, two sources of individual differences can be recognised. First, a difference in sensitivity to the prosodic manipulation (variability along the solid line). Second, a difference in response bias (variability along the dashed line). In the syntactic decisions (Fig 8A) direction of this bias differed strongly

between participants, showing some participants with an overall bias towards nominative-determiner responses (above the solid line), and others towards accusative-determiner responses (below the solid line), regardless of the focus position (see Fig 8A). In the semantic decisions, a bias for agent-like nouns was visible (most participants above the solid line), and the range in bias was less wide than in the syntactic decisions (Fig 8B).

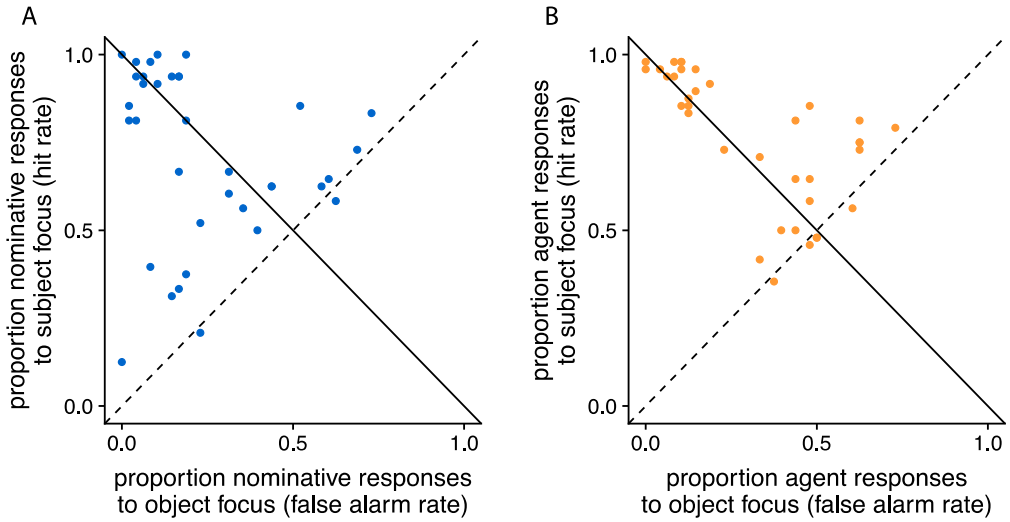



**Figure 8**: Signal detection theory analysis for the syntactic (A) and semantic decisions (B) in Experiment 2. Dots represent individual subjects. Positions distanced further above the dashed line indicate a higher sensitivity. Positions on either side of the solid line indicate a response bias towards a subject-focus (right side) or an object-focus interpretation (left side).

## 4. Discussion

This study shows that in online sentence comprehension pitch accents establish dissociable linguistic expectations about upcoming sentence elements. Since pitch accents can influence the interpretation of a sentence by marking contrasts and parallels between constituents of a sentence (Carlson, 2001; 2015; Carlson et al., 2009), we hypothesised that pitch accents establish expectations about syntactic and semantic aspects of the upcoming constituents. To test the existence of these expectations and whether they can be probed implicitly or explicitly, we used sentences with contrastive focus and an ellipsis structure in two experiments, using a sentence comprehension and a sentence completion paradigm. The results show that pitch accents, by highlighting constituents that contain syntactic and semantic cues, establish expectations concerning both the syntactic and semantic properties of an upcoming noun phrase. Results of the sentence comprehension task (Experiment 1) revealed that participants built syntactic expectations implicitly: when pitch accents marked syntactic information that was met with mismatching syntactic information later in the sentence, responses were slower. This effect was not found for mismatching semantic information. In turn, the sentence completion task (Experiment 2) provided evidence for both syntactic and semantic expectations, when explicitly probed. Participants were able to complete the sentence with a determiner and noun in agreement with the respective syntactic and semantic properties of the pitch-accented noun phrase in the preceding clause. Finally, our results demonstrate that, when contradictory cues occur in the same sentence, the syntactic cue (case-marking) takes precedence over the semantic cue (thematic role), and previous prosodically-cued information is overwritten. These data reveal that during auditory sentence comprehension prosodic, semantic and syntactic information types are processed to create expectations about the upcoming linguistic elements in the sentence. All information types are used online, but there is a clear precedence for local, unambiguous syntactic information when assigning a constituent's role in the sentence.

The first experiment showed that pitch accents form a syntactic expectation about the upcoming sentence structure. This expectation could be probed *implicitly* by measuring the effects that contradictory syntactic cues in noun phrases receiving contrastive pitch accents have on sentence comprehension. Contradictory syntactic cues in the two focused constituents led to lengthened reaction times when the participants were asked about the

role of the first constituent (i.e. *What was the role of the policeman/thief?*). Importantly, in these trials, where only the interpretation of the first part of the sentence was tested, the mismatch was irrelevant for the task: the probed noun phrase in the main clause always held congruent syntactic-semantic information and its interpretation was independent of the cues provided in the ellipsis, where the violation occurred. This finding supports the notion that the two constituents in contrastive focus are interpreted to fulfil parallel roles, occupying syntactically identical functions (Carlson, 2001; Stolterfoht et al., 2007; Winkler, 2019). Furthermore, it suggests that after having heard the pitch accent on the first contrasted noun phrase, a specific expectation about the syntactic properties of the second contrasted noun phrase is formed. A violation of this assumption (by mismatching grammatical case) results in additional processing costs. This finding provides further insight into the role of pitch accents in the prosody-syntax interface. In addition to disambiguating case (Weber, Grice, & Crocker, 2006b) and resolving structural (Nakamura et al., 2012) and attachment ambiguities (Carlson & Tyler, 2017), pitch accents can also establish expectations about the syntactic role of upcoming sentence elements early on in the sentence. We suggest that, in case of an ellipsis site without any structural information (e.g., in remnants with a proper name, such as in the earlier example "…, not Peter"), this syntactic expectation influences the interpretation of the ellipsis structure. While we hypothesised that contradictory semantic cues could have analogous effects on sentence processing (e.g., longer reaction times), the results of Experiment 1 did not show evidence that focus had established semantic expectations. We will discuss these differences between the processing of syntactic and semantic cues below, but will first address the response behaviour at the ellipsis site.

The response behaviour at the ellipsis site (Experiment 1) allows for several conclusions concerning the relative dominance of the prosodic, syntactic, and semantic cues when contradictory information is present. First, the syntactic expectation established by focus-marking in the main clause was not strong enough to interfere with the interpretation of the ellipsis noun phrase. Rather, when this noun phrase was probed (*What was the role of the inspector/murderer?*), participants based their response on the local syntactic cue (their subject/object interpretation followed the case of the determiner). This suggests that the local syntactic cue had overwritten the syntactic expectation established by focus-marking in

the main clause. The second observation is that the local semantic cue did influence the response at the ellipsis site. We found slower responses to *the*$^{NOM}$ *MURDERER* as compared to *the*$^{ACC}$ *INSPECTOR* as well as a significant decrease in the number of "subject" responses. In accordance with a large body of research showing that thematic role typicality can influence syntactic parsing (Trueswell et al., 1994), this suggests that the thematic content of the object-typical noun MURDERER cued a syntactic role that was incompatible with the preceding syntactic cue of the determiner (*the*$^{NOM}$, assigning subject role), yielding an interpretation that was difficult to process. The finding that the subject-interpretation of *the*$^{NOM}$ *MURDERER* led to processing difficulties whereas the object-interpretation of *the*$^{ACC}$ *INSPECTOR* (with *the*$^{ACC}$ assigning object role) did not, may be due to the type of verb-argument items used in our stimulus set. In most items, it was less plausible for the noun phrases in patient-role to reverse their typical role (i.e., to adopt an agent-role) than vice versa. Yet, regardless of this semantic effect, we can conclude that, in the type of construction investigated in our study, the syntactic cue was decisive for the interpretation of the ellipsis. This may also explain why in the responses to the constituents in the main clause, effects of the syntactic violation were stronger than those of the semantic violation, since a more decisive cue may lead to more disruptive processing once violated.

The stronger reaction time effect observed in the syntactic violations as compared to the semantic violations in responses to the main clause can be explained in several other ways. A first, straightforward explanation may be given by the different nature of the two cues: grammatical case is invariably mapped to subject and object roles of a sentence, whereas the thematic role of a noun is dependent on the semantic features of the verb and accompanying arguments. It is plausible to assume that in the sentence construction under investigation, the syntactic cue of the noun phrase highlighted by focus is more decisive in establishing the sentence structure, because its binary nature (nominative/accusative) makes it more categorical than the semantic cue. Alternatively, a general lack of reliability of the semantic expectation during the whole experiment may have diminished the relative effect of these cues, since our sentences contained a semantic conflict in approximately half of the trials: the semantic mismatch between the two focused constituents in the semantics-focus mismatch condition, in addition to the local thematic incongruency at the ellipsis site (*the*$^{NOM}$ *murderer* and *the*$^{ACC}$ *inspector)*. Indeed, if predictions are disconfirmed frequently

enough in an experiment, their predictive strength is diminished (Brothers, Swaab, & Traxler, 2017). Similarly, intonational cues have recently been suggested to lose their predictive value when the listener deems them unreliable (Roettger & Franke, 2019), which, in combination with the semantic expectations possibly being weaker than the syntactic expectations, may have contributed to the lack of effect in the semantics-focus mismatch condition.

An exploratory analysis of our effects over time revealed that there might have been a small effect of the semantics-focus mismatch at the beginning of the experimental session. These visualisations of the responses over time also suggest that over the course of the experiment listeners started disregarding the semantic and prosodic cues. Diminished processing of the prosodic cues would have decreased the syntactic mismatch established by focus, which would explain why the effect in the reaction times of the syntax-focus mismatch condition was stronger at the beginning of the experiment. Future studies would be necessary to further investigate the time-dependency of these effects and the adoption of explicit strategies, for example by using an increased number of catch trials (this may prevent listeners from weighing cues differently over time). A limitation to the current design is the relatively small number of catch trials, however, including more catch trials was not feasible: the complexity of our experimental design (including several experimental factors and multiple levels within each factor) required a large number of trials to achieve sufficient statistical power, resulting in a long running time of the experiment. Finally, it has to be considered that semantic expectations might not have been formed at all: from Experiment 1, we could not conclude if the expectations formed by pitch accents were too weak to lead to lengthened response times when violated, or if they had not been established in the first place.

Since Experiment 1 did not yield conclusive evidence concerning expectations established in the semantic domain, we aimed to probe these expectations directly using the same material in an alternative paradigm in Experiment 2. Our stimulus design enabled us to assess if focus-marking establishes *explicit* expectations, by employing a sentence completion task that teased apart syntactic and semantic decisions. Here, we found evidence for prosodically-formed expectations in both the syntactic and semantic domains.

The results of Experiment 2 showed that participants preferred to complete the sentence with syntactic and semantic elements that corresponded to the focused constituent in the main clause. After subject focus, participants preferred to continue the sentence with a determiner in the nominative case and an agent-like noun. The opposite pattern was found after object-focused sentences. This evidence of syntactic expectations is in line with our result of Experiment 1. In addition, Experiment 2 showed that focus can in fact establish a semantic expectation: participants showed a preference for typical-agent nouns after subject focus sentences and typical patient-nouns after object focus. Since focus activates a set of alternatives to the focused noun (Gotzner & Spalek, 2019), it is likely that listeners activated nouns associated with thematic roles to the verb. Our results show that depending on the focus location, they subsequently selected a noun associated with either the subject- or object-role of the pitch-accented constituent. However, it remains to be explained why semantic expectations were revealed when explicitly probed, yet did not cause an increase in response times in the main clause, semantics-focus mismatch trials of Experiment 1. This discrepancy between the two experiments supports the idea that, in contrast to the syntactic cues, the semantic cues were not decisive enough to lead to processing costs when violated. Indeed, the effects of semantic cues in establishing parallels between constituents have been shown to be relatively small (Carlson, 2015; Carlson et al., 2009). Furthermore, discrepancies between results from offline tasks such as sentence completion and online tasks (EEG or eye-tracking) have been reported previously (Chow, Smith, Lau, & Phillips, 2015; Karimi et al., 2019). Considering these task-dependent differences, our results may suggest that, even though semantic expectations could be established by pitch accents and subsequently accessed during offline processing, their role in online processing is not decisive enough to yield measurable effects.

Notably, while focus-marking influenced syntactic and semantic responses in Experiment 2, some listeners responded more according to their inherent biases in both domains (see Figures 8A and B). Previous studies in both German (Stolterfoht et al., 2007) and English (Carlson et al., 2009) have provided evidence for the existence of a default interpretation concerning the information structure of a sentence such as “Yesterday the policeman arrested the thief, not the murderer”. Listeners tend to show a bias to assign prominence late in the sentence, to the object noun phrase (*the thief*). These studies

showed that prosodic (Carlson, 2015) and semantic factors (Carlson, 2001) have limited effects in shifting this interpretation, and the inherent bias usually persists. However, most of the previous studies used grammaticality judgements, questionnaires, or (self-paced) reading. By explicitly probing the preferred syntactic or semantic structure of the upcoming phrase, we were able to obtain a direct measure of the perceived focus position and a possible bias. The bias in the semantic responses indicated a response preference for agent-like nouns, that is, significantly fewer focus-congruent responses after object vs subject focus trials. This may be explained by differences between processing subject and object roles more generally. One of the most consistent properties of case systems across languages is the prominent role attributed to the subject that is simultaneously the agent of a transitive verb (Bickel, Witzlack-Makarevich, Choudhary, Schlesewsky, & Bornkessel-Schlesewsky, 2015). As a result, the nominative marking of a noun phrase is particularly salient and triggers a strong mismatch response if stereotypical semantic features of a subject are not met (widely investigated as semantic reversal anomalies (Bornkessel-Schlesewsky et al., 2011)). Altogether, the subject marking presupposes a typical, narrower profile with particular features, whereas object marking usually does not trigger such expectations. This may explain our finding in the ellipsis trials of Experiment 1, showing that the conflicting determiner-noun pair *the*$^{NOM}$ *+ typical patient* led to slower responses than *the*$^{ACC}$ *+ typical agent*. Likewise, this reasoning could explain the response pattern in Experiment 2, showing an advantage for agent-like nouns in the continuation of subject-focus sentences over patient-like nouns in an object-focus setting.

Participants also showed a bias in their syntactic responses: the response proportions of Experiment 2 indicated that participants had a preference for accusative determiners, confirming the bias for an object-focus interpretation reported previously (Stolterfoht et al., 2007). These results are in line with the differences in reaction times from Experiment 1, where exploratory analyses suggested a preference for object focus sentences when the comprehension question probed the ellipsis. Conversely, in the main clause, performance indicated a preference for the subject focus trials. This latter effect, showing a subject-focus preference when the main clause was probed, may be explained by the fact that prosodically, postfocal words are usually deaccented (Féry & Kügler, 2008): perhaps the subject focus construction was therefore perceived as being more salient, possibly yielding

additional attentional and memory effects. Alternatively, focus projected by a pitch accent on the subject noun phrase has been shown to be perceived as narrower as compared to focus projected by a pitch accent on the object (Kuthy & Stolterfoht, 2019). The narrower interpretation of subject focus may have led to an advantage over processing object focus sentences in the main clause. Finally, the fact that the behaviour in Experiment 1 shows subject vs. object effects in opposite directions in the main clause as compared to the ellipsis trials may explain why in our online norming study, no difference was found between the two focus structures of the baseline condition (see Section 2.1.2): the norming study probed the interpretation of the baseline sentences as a whole, and the opposite biases for main clause and ellipsis responses may have cancelled out.

We observed considerable inter-individual variability within the syntactic and semantic biases in Experiment 2. Inter-individual variability in syntactic attachment is a well-known phenomenon and has previously been linked to differences in working-memory constraints (Swets, Desmet, Hambrick, & Ferreira, 2007). Such variability has also been reported in prosody processing, both in the perception (Roy, Cole, & Mahrt, 2017) and production of prosodic cues (Ferreira & Karimi, 2015; Xie, Buxó-Lugo, & Kurumada, 2021), as well as in implicit prosody perception (Jun & Bishop, 2015). An important observation that can be made in the present results is that some listeners appear to rely on their biases, whereas others rely more strongly on the prosodic signal. This result is in line with a recent study showing inter-individual variability in the acoustic and linguistic variables used by listeners to determine prominence (Baumann & Winter, 2018). A worthwhile avenue for future research would be to further investigate the factors that determine whether a listener is rather led by acoustic cues or inherent bias in perceiving prosodic events. Finally, for the listeners that responded according to an inherent bias in our study, the results do not allow us to determine whether the source of that bias was at the perceptual or at the response level. One possibility is that listeners had a *perceptual* bias for either subject or object focus constructions. Alternatively, the participants could possess a *response* bias for a specific syntactic or semantic structure at the ellipsis site. Further research is required to tease these two explanations apart.

In conclusion, this study sheds new light on the interfaces of prosody with syntactic structure and with information structure. We show that pitch accents can establish expectations on upcoming sentence elements. Here, separate syntactic and semantic processes can be distinguished and only the expectations in the syntactic domain were decisive enough to increase processing costs when violated. Furthermore, our design enabled us to draw conclusions concerning the relative dominance of syntactic, semantic, and prosodic cues in guiding sentence comprehension. In case of multiple contradictory cues, we show that the effects of prosodically cued expectations are limited and readily overwritten by local syntactic cues. This is in line with the notion that the role of prosody in sentence comprehension is influential, but not decisive (e.g. Frazier, Carlson, & Clifton, 2006; Carlson, 2009). Finally, we could observe individual differences within the use of pitch accents in establishing sentence structure, and we put forward that future studies should further investigate the factors that make a listener rely on bottom-up acoustic information or rather be driven by top-down internal biases.

**Acknowledgements**

We wish to thank Isabel Gebhardt, Lisa Kunz, and Lisa Reimund for help with data acquisition, Heike Boethel for organisational support, and all volunteers for participation. Furthermore, we thank Jana Klaus and Vincent K. M. Cheung for analysis scripts and insights and Thomas C. Gunter, Emiliano Zaccarella, and colleagues of the Neuropsychology department and Lise Meitner research group for stimulating discussions. Special thanks go to Giorgio Papitto for help in setting up the online rating experiment. Finally, we would like to thank our 3 anonymous reviewers for their careful and thorough reviews. The study was funded by the Max Planck Society.

**Conflict of interest**

The authors declare that there are no conflicts of interest.

## References


Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. (2018). Raincloud plots: a multi-platform tool for robust data visualization. http://doi.org/10.7287/peerj.preprints.27137v1

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*(4), 390–412. http://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. http://doi.org/10.1016/j.jml.2012.11.001

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). Retrieved from https://philpapers.org/rec/BATFAT-3

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). http://doi.org/10.18637/jss.v067.i01

Baumann, S., & Winter, B. (2018). What makes a word prominent? Predicting untrained German listeners' perceptual judgments. *Journal of Phonetics*, *70*, 20–38. http://doi.org/10.1016/j.wocn.2018.05.004

Bickel, B., Witzlack-Makarevich, A., Choudhary, K. K., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2015). The Neurophysiology of Language Processing Shapes the Evolution of Grammar: Evidence from Case Marking. *PloS One*, *10*(8), e0132819. http://doi.org/10.1371/journal.pone.0132819

Boersma, P., & Weenink, D. (n.d.). Praat. Retrieved from http://www.praat.org/

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., et al. (2011). Think globally: cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language*, *117*(3), 133–152. http://doi.org/10.1016/j.bandl.2010.09.010

Braun, B., & Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*.

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, *93*, 203–216. http://doi.org/10.1016/j.jml.2016.10.002

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), 1052. http://doi.org/10.5334/joc.10

Carlson, K. (2001). The Effects of Parallelism and Prosody in the Processing of Gapping Structures. *Language and Speech*, *44*(1), 1–26. http://doi.org/10.1177/00238309010440010101

Carlson, K. (2002). Parallelism and prosody in the processing of ellipsis sentences. New York; London: Routledge.

Carlson, K. (2009). How Prosody Influences Sentence Comprehension. *Language and Linguistics Compass*, *3*(5), 1188–1200. http://doi.org/10.1111/j.1749-818X.2009.00150.x

Carlson, K. (2015). Clefting, Parallelism, and Focus in Ellipsis Sentences. In K. Carlson, C. Clifton, & J. D. Fodor (Eds.), *Grammatical Approaches to Language Processing* (2nd ed., Vol. 46, pp. 63–83). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-12961-7_4

Carlson, K., & Tyler, J. C. (2017). Accents, Not Just Prosodic Boundaries, Influence Syntactic Attachment. *Language and Speech*, 23830917712282. http://doi.org/10.1177/0023830917712282

Carlson, K., Dickey, M. W., Frazier, L., & Clifton, C., Jr. (2009). Information structure expectations in sentence comprehension. *The Quarterly Journal of Experimental Psychology*, *62*(1), 114–139. http://doi.org/10.1080/17470210701880171

Chow, W.-Y., Smith, C., Lau, E., & Phillips, C. (2015). A "bag-of-arguments" mechanism for initial verb predictions. *Language, Cognition and Neuroscience*, *31*(5), 577–596. http://doi.org/10.1080/23273798.2015.1066832

Clahsen, H., & Featherston, S. (1999). Antecedent priming at trace positions: Evidence from German scrambling. *Journal of Psycholinguistic Research*. http://doi.org/10.1023/A:1023293132656

Culicover, P. W. & Jackendoff, R. *Simpler Syntax*. (Oxford University Press, 2005).

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development*, *79*(4), 1152–1167. http://doi.org/10.1111/j.1467-8624.2008.01181.x

Dobson, A. J., & Barnett, A. G. (2008). An introduction to generalized linear models. Boca Raton, FL: CRC Press.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*(2), 164–203. http://doi.org/10.1016/S0010-0285(03)00005-7

Ferreira, F., & Karimi, H. (2015). Prosody, Performance, and Cognitive Skill: Evidence from Individual Differences. In L. Frazier & E. Gibson (Eds.), *Grammatical Approaches to Language Processing* (Vol. 46, pp. 119–132). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-12961-7_7

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, *36*(4), 680–703. http://doi.org/10.1016/j.wocn.2008.05.001

Frazier, L., Carlson, K., & Clifton, C., Jr. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249. http://doi.org/10.1016/j.tics.2006.04.002

Ginzburg, J. & Sag, I. A. *Interrogative Investigations*. (Stanford Univ Center for the Study of Language and Information, 2000). doi:10.1145/3373625.3418047

Gotzner, N., & Spalek, K. (2019). The life and times of focus alternatives: Tracing the activation of alternatives to a focused constituent in language comprehension. *Language and Linguistics Compass*, *13*(2), e12310. http://doi.org/10.1111/lnc3.12310

Grabe, E. (1998). Pitch accent realization in English and German. *Journal of Phonetics*, *26*(2), 129–143. http://doi.org/10.1006/jpho.1997.0072

Husband, E. M., & Ferreira, F. (2016). The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience*, *31*(2), 217–235.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*(2), 541–573. http://doi.org/10.1016/j.jml.2007.06.013

Jackendoff, R. S. (1972). Semantic Interpretation in Generative Grammar. The MIT Press, 28 Carleton Street, Cambridge, Mass. 02142 ($10.00).

Jun, S.-A., & Bishop, J. (2015). Priming Implicit Prosody: Prosodic Boundaries and Individual Differences, 1–15. http://doi.org/10.1177/0023830914563368

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of Syntactic and Semantic Information in Predictive Processing: Cross-Linguistic Evidence from German and English. *Journal of Psycholinguistic Research*, *32*(1), 37–55. http://doi.org/10.1023/A:1021933015362

Karimi, H., Brothers, T., & Ferreira, F. (2019). Phonological versus semantic prediction in focus and repair constructions: No evidence for differential predictions. *Cognitive Psychology*, *112*, 25–47. http://doi.org/10.1016/j.cogpsych.2019.04.001

Kuthy, K. D., & Stolterfoht, B. (2019). Focus projection revisited: Pitch accent perception in German. In S. Featherston, R. Hörnig, S. V. Wietersheim, & S. Winkler (Eds.), *Experiments in Focus* (pp. 57–70). Berlin, Boston: De Gruyter. http://doi.org/10.1515/9783110623093-003

Lee, E.-K., & Watson, D. G. (2010). Effects of pitch accents in attachment ambiguity resolution. *Language and Cognitive Processes*, *26*(2), 262–297. http://doi.org/10.1080/01690965.2010.491650

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (n.d.). emmeans: Estimated Marginal Means, aka Least-Squares Means.

Macmillan, N. A., & Creelman, C. D. (1991). Detection Theory. Cambridge University Press.

Merchant, J. *The Syntax of Silence*. (Oxford University Press on Demand, 2001).

Merchant, J. in *Oxford Handbook of Ellipsis* (eds. van Craenenbroeck, J. & Temmerman, T.) (2018). doi:10.1093/oxfordhb/9780198712398.013.2

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N., & Friederici, A. D. (2016). Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cerebral Cortex*. http://doi.org/10.1093/cercor/bhw228

Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition*, *125*(2), 317–323. http://doi.org/10.1016/j.cognition.2012.07.016

Oldfield, R. C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* **9,** 97–113 (1971).

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., et al. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. http://doi.org/10.3758/s13428-018-01193-y

Roettger, T. B., & Franke, M. (2019). Evidential Strength of Intonational Cues and Rational Adaptation to (Un-)Reliable Intonation. *Cognitive Science*, *43*(7), 1. http://doi.org/10.1111/cogs.12745

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, *1*(1), 75–116. http://doi.org/10.1007/BF02342617

Ross, J. R. Guess Who? in (eds. Binnick, R. I., Davison, A., Green, G. M. & Morgan, J. L.) 252–286 (1969).

Roy, J., Cole, J., & Mahrt, T. (2017). Individual differences and patterns of convergence in prosody perception. *Laboratory Phonology*, *8*(1), 22. http://doi.org/10.5334/labphon.108

Schafer, A. J., Carlson, K., Clifton, C., & Frazier, L. (2000). Focus and the Interpretation of Pitch Accent: Disambiguating Embedded Questions:. *Language and Speech*, *43*(1), 75–105. http://doi.org/10.1177/00238309000430010301

Schafer, A. J., Carter, J., Clifton, C., Jr, & Frazier, L. (1996). Focus in relative clause construal. *Language and Cognitive Processes*.

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. *The American Statistician*, *34*(4), 216–221. http://doi.org/10.1080/00031305.1980.10483031

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. H. Spieler & E. Schumacher (Eds.), *New Methods in Neuroscience and Cognitive Psychology* (pp. 4–31).

Snedeker, J., & Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 1–28.

Stolterfoht, B. (2005) Processing word order variations and ellipses the interplay of syntax and information structure during sentence comprehension. (MPI for Human Cognitive and Brain Sciences).

Stolterfoht, B., Friederici, A. D., Alter, K., & Steube, A. (2007). Processing focus structure and implicit prosody during reading: differential ERP effects. *Cognition*, *104*(3), 565–590. http://doi.org/10.1016/j.cognition.2006.08.001

Strotseva-Feinschmidt, A., Cunitz, K., Friederici, A. D., & Gunter, T. C. (2015). Auditory Discrimination Between Function Words in Children and Adults: A Mismatch Negativity Study. *Frontiers in Psychology*, *6*(213), e22999. http://doi.org/10.3389/fpsyg.2015.01930

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: a psychometric approach. *Journal of Experimental Psychology: General*, *136*(1), 64–81. http://doi.org/10.1037/0096-3445.136.1.64

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic Influences On Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution. *Journal of Memory and Language*, *33*(3), 285–318. http://doi.org/10.1006/jmla.1994.1014

Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting Pitch Accents in Online Comprehension: H* vs. L+H*. *Cognitive Science*, *32*(7), 1232–1244. http://doi.org/10.1080/03640210802138755

Weber, A., Braun, B., & Crocker, M. W. (2006a). Finding referents in time: eye-tracking evidence for the role of contrastive accents. *Language and Speech*, *49*(Pt 3), 367–392. http://doi.org/10.1177/00238309060490030301

Weber, A., Grice, M., & Crocker, M. W. (2006b). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition*, *99*(2), B63–72. http://doi.org/10.1016/j.cognition.2005.07.001

Winkler, S. (2019). Ellipsis and Prosody. In J. van Craenenbroeck & T. Temmerman (Eds.), *Oxford Handbook of Ellipsis* (pp. 356–386). Oxford University Press. http://doi.org/10.1093/oxfordhb/9780198712398.013.15

Xie, X., Buxó-Lugo, A., & Kurumada, C. (2021). Encoding and decoding of meaning through structured variability in intonational speech prosody. *Cognition, 211*, Article 104619. https://doi.org/10.1016/j.cognition.2021.104619.

## Appendix

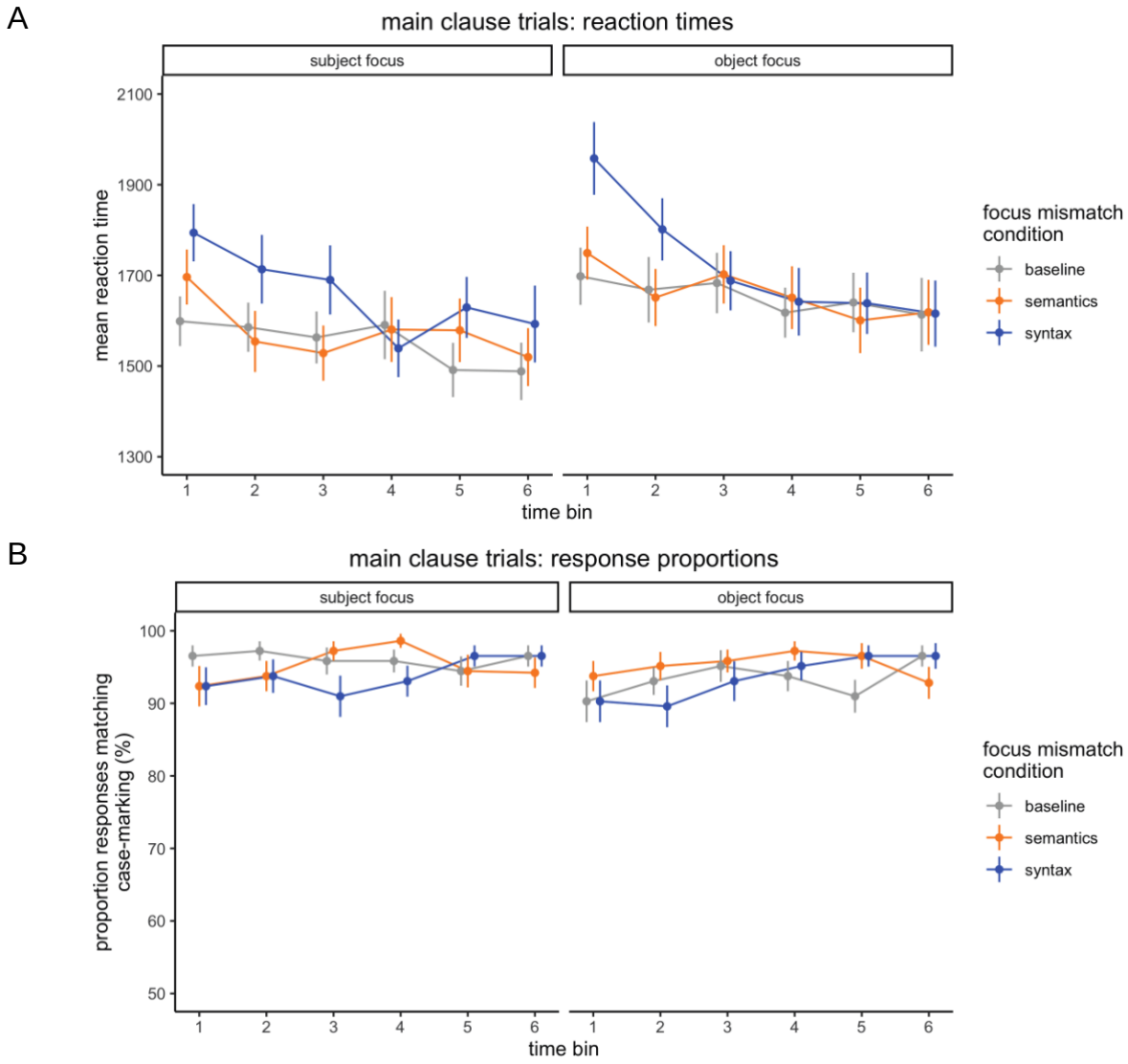



***Figure A1:*** *Exploratory visualisation of reaction times (A) and response proportions (B) of the main clause trials in Experiment 1. Results are plotted in function of focus mismatch condition and focus position, split across six time bins. Error bars indicate ±1 SEM.*

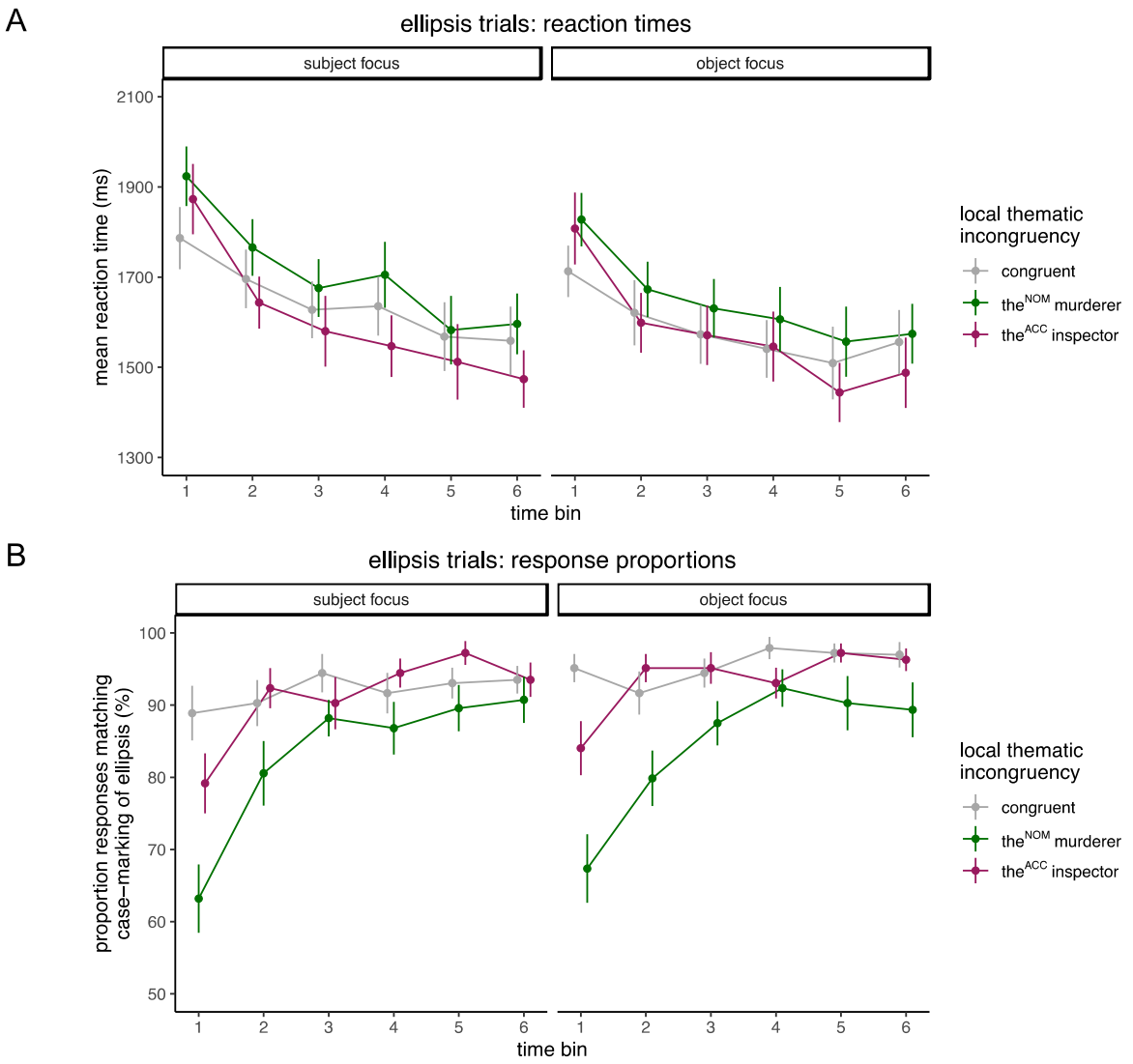



***Figure A2:*** *Exploratory visualisation of reaction times (A) and response proportions (B) of the ellipsis trials in Experiment 1. Results are plotted in function of focus mismatch condition and focus position, split across six time bins. Error bars indicate ±1 SEM.*

***Table A1:*** *Analysis of log response times (Experiment 1). Results from linear mixed effects model including the fixed effects violation type (baseline, semantic, syntactic), focus position (subject, object) and comprehension question target (main clause, ellipsis). Random effects included intercepts for participants and items. Model formula: log(RT) ~ violation type * focus position * comprehension question target + (1 | subject) + (1 | verb)*

| ***fixed effect*** | | ***estimate*** | ***SE*** | ***t*** |
|---|---|---|---|---|
| intercept | | 7.343 | 0.032 | 226.182 |
| violBaseline | | -0.012 | 0.006 | -1.967 |
| violSyntactic | | 0.016 | 0.006 | 2.527 |
| probeEllipsis | | -0.002 | 0.004 | -0.377 |
| focObject | | 0.008 | 0.004 | 1.881 |
| violBaseline * probeEllipsis | | 0.002 | 0.006 | 0.314 |
| violSyntactic * probeEllipsis | | 0.025 | 0.006 | 3.981 |
| violBaseline * focObject | | -0.013 | 0.006 | -2.037 |
| violSyntactic * focObject | | -0.027 | 0.006 | -4.269 |
| probeEllipsis * focObject | | -0.035 | 0.004 | -7.961 |
| violBaseline * probeEllipsis * focObject | | 0.010 | 0.006 | 1.643 |
| violSyntactic * probeEllipsis * focObject | | 0.034 | 0.006 | 5.399 |
| ***random effect*** | | | ***variance*** | ***SD*** |
| verb | intercept | | 0.001 | 0.034 |
| subj | intercept | | 0.036 | 0.191 |
| residual | | | 0.067 | 0.259 |

***Table A2:*** *Analysis of response proportions (Experiment 1). Results from generalized linear mixed effects model including the fixed effects violation type (baseline, semantic, syntactic), focus position (subject, object) and comprehension question target (main clause, ellipsis). Random effects included intercepts for participants and items, and by-participant slopes for the factor violation type. Model formula: response ~ violation type * focus position * comprehension question target + (1 + violation type | subject) + (1 | verb)*

| ***fixed effect*** | | | ***estimate*** | ***SE*** | ***z*** | ***p*** |
|---|---|---|---|---|---|---|
| intercept | | | 2.788 | 0.152 | 18.290 | <.001 |
| violBaseline | | | 0.284 | 0.133 | 2.130 | 0.033 |
| violSyntactic | | | -0.175 | 0.113 | -1.543 | 0.123 |
| focObject | | | -0.285 | 0.070 | -4.061 | <.001 |
| probeEllipsis | | | 0.447 | 0.070 | 6.348 | <.001 |
| violBaseline * focObject | | | 0.269 | 0.104 | 2.594 | 0.009 |
| violSyntactic * focObject | | | 0.472 | 0.095 | 4.987 | <.001 |
| violBaseline * probeEllipsis | | | -0.369 | 0.104 | -3.563 | <.001 |
| violSyntactic * probeEllipsis | | | -0.088 | 0.095 | -0.934 | 0.350 |
| focObject * probeEllipsis | | | 0.277 | 0.070 | 3.937 | <.001 |
| violBaseline * focObject * probeEllipsis | | | 0.032 | 0.104 | 0.306 | 0.759 |
| violSyntactic * focobject * probeEllipsis | | | -0.436 | 0.095 | -4.585 | <.001 |
| ***random effect*** | | | | | ***variance*** | ***SD*** |
| verb | intercept | | | | 0.031 | 0.177 |
| subj | intercept | | | | 0.565 | 0.752 |
| | violBaseline | | | | 0.090 | 0.300 |
| | violSyntactic | | | | 0.018 | 0.135 |
| | intercept | violBaseline | | | -0.057 | -0.251 |
| | intercept | violSyntactic | | | 0.030 | 0.299 |
| | violBaseline | violSyntactic | | | -0.021 | -0.509 |

***Table A3:*** *Analysis of log response times (Experiment 2). Results from linear mixed effects model including the fixed effects decision type (semantic, syntactic) and focus position (subject, object). Random effects included intercepts for participants and items, and by-participant and by-item slopes for the interaction decision type x focus position. Model formula: log(RT) = decision type * focus position + (1 + decision type + violation type | subject) + (1 | verb)*

| ***fixed effect*** | | ***estimate*** | ***SE*** | ***t*** |
|---|---|---|---|---|
| intercept | | 6.861 | 0.036 | 190.693 |
| decisionSyntactic | | -0.131 | 0.016 | -8.033 |
| focusObject | | -0.009 | 0.005 | -1.829 |
| decisionSyntactic * focusObject | | 0.018 | 0.004 | 4.607 |
| ***random effect*** | | | ***variance*** | ***SD*** |
| verb | intercept | | 0.001 | 0.037 |
| subj | intercept | | 0.045 | 0.212 |
| | decisionSyntactic | | 0.009 | 0.095 |
| | focusObject | | 0 | 0.015 |
| | intercept, decisionSyntactic | | 0.005 | 0.246 |
| | intercept, focusObject | | -0.001 | -0.410 |
| | decisionSyntactic, focusObject | | 0 | -0.327 |
| residual | | | 0.106 | 0.326 |

***Table A4****: Analysis of response proportions (Experiment 2). Results from generalized linear mixed effects model including the fixed effects decision type (semantic, syntactic) and focus position (subject, object). Random effects included intercepts for participants and for items. Model formula: response ~ decision type * focus position + (1 | subj) + (1 | verb)*

| ***fixed effect*** | | ***estimate*** | ***SE*** | ***z*** | ***p*** |
|---|---|---|---|---|---|
| intercept | | 1.307 | 0.200 | 6.545 | <.001 |
| decisionSyntactic | | 0.030 | 0.030 | 1.029 | 0.303 |
| focusObject | | 0.036 | 0.030 | 1.217 | 0.224 |
| decisionSyntactic * focusObject | | -0.187 | 0.030 | -6.326 | <.001 |
| ***random effect*** | | | ***variance*** | ***SD*** | |
| verb | intercept | | 0.012 | 0.108 | |
| subject | | intercept | 1.376 | 1.173 | |