Kyriaki, Louise & Schlesewsky, Matthias & Bornkessel-Schlesewsky, Ina. 2021. Thematic role assignment in non-default verb classes: A cross-linguistic comparison of English and German. *Glossa: a journal of general linguistics* 6(1): 134, pp. 1–26. DOI: https://doi.org/10.16995/glossa.5728

Thematic role assignment in non-default verb classes: A cross-linguistic comparison of English and German

Louise Kyriaki, Cognitive and Systems Neuroscience Research Hub, GPO Box 2471, Adelaide SA 5001, Australia, Louise.Kyriaki@mymail.unisa.edu.au

Matthias Schlesewsky, Cognitive and Systems Neuroscience Research Hub, GPO Box 2471, Adelaide SA 5001, Australia

Ina Bornkessel-Schlesewsky, Cognitive and Systems Neuroscience Research Hub, GPO Box 2471, Adelaide SA 5001, Australia

The influence of sentential cues (such as animacy and word order) on thematic role interpretation differs as a function of language (MacWhinney et al. 1984). However, existing cross-linguistic research has typically focused on transitive sentences involving agents, and interpretation of non-default verb classes is less well understood. Here, we compared the way in which English and German native speakers – languages known to differ in the cue prominence of animacy and word order – assign thematic roles. We compared their interpretation of sentences containing either default (agent-subject) or non-default (experiencer-subject) verb classes. Animacy of the two noun phrases in a sentence was either animate-inanimate and plausible (e.g. "The men will devour the meals…") or inanimate-animate and implausible in English (e.g. "The meals will devour the men…"). We examined role assignment by probing for either the actor or undergoer of the sentence. Mixed effects modelling revealed that role assignment was significantly influenced by noun animacy, verb class, question type, and language. Results are interpreted within the Competition Model framework (Bates et al. 1982; MacWhinney et al. 1984) and show that English speakers predominantly relied on word order for thematic role assignment. German speakers relied on word order to a comparatively lesser degree, with animacy a prominent cue. Cue weightings appeared to be modulated in the context of other cues, with the weighting of an animacy-based strategy over a word-order-based strategy increasing for sentences with non-default (experiencer-subject) verbs and with undergoer-focused questions, particularly where word order was more flexible (i.e. in German as opposed to English). These findings highlight the differential influence of the surrounding context (e.g. question focus) across languages.

1. Introduction

Psycholinguistic studies of cross-linguistic sentence comprehension have predominantly examined a narrow range of sentence constructions. While this allows for effective comparison between languages, current findings are almost exclusively focused on default verb classes (agent-subject, default case marking). For example, seminal crosslinguistic work within the context of the Competition Model (CM; Bates et al. 1982) involved presenting native speakers with simple transitive sentences containing two concrete nouns and a transitive action verb, followed by a prompt to identify the actor/subject (Bates & MacWhinney 1989). Such experiments allowed for the classification of the syntactic and semantic cues that enable speakers to identify thematic roles, including word order, agreement, animacy, and stress (MacWhinney et al. 1984). According to the CM, the relative use of these cues differs across languages depending on cue availability, reliability, and complexity (MacWhinney et al. 1984).

This literature has revealed that for simple transitive sentences with default verb classes, many languages have a dominant cue, which – when present – (almost) deterministically guides interpretation. For example, while English predominantly relies on word order, German sentence interpretation is dominated by case marking (when unambiguously present) and Italian draws primarily on subject-verb agreement (MacWhinney et al. 1984). Recent neurolinguistic work also shows qualitative differences in the brain's comprehension strategy depending on the dominant cue of the language being comprehended (Bornkessel-Schlesewsky et al. 2011; Bornkessel-Schlesewsky & Schlesewsky 2020).

A growing body of literature has investigated verb classes which deviate substantially from default configurations, such as experiencer-object verbs which call for an inversion of the thematic hierarchy (Experiencer (object) > Stimulus (subject)). Studies examining sentence comprehension with these verb types – which are sometimes characterised as unaccusative following Belletti & Rizzi (1988) – suggest that cue use differs in object-experiencer constructions compared to default configurations. For example, the preference for subject-before-object (nominative before non-nominative) orders appears to be weaker in sentences with object-experiencer verbs in both German (Schlesewsky & Bornkessel 2003; Bornkessel et al. 2004) and Spanish (Gattei et al. 2015; 2018). Thus, at least in languages that allow some degree of word order flexibility, the strength of the word order cue is bolstered when multiple hierarchies are aligned (thematic, grammatical function, case); conversely, it decreases in the case of a hierarchy mismatch (for related observations from language acquisition, see Primus & Lindner 1994).

However, it is unclear how comprehension is affected in constructions where the deviation from the default is subtler, for example with experiencer-subject verbs. Here, the hierarchy between thematic roles and grammatical functions aligns (Experiencer (subject) > Stimulus/Theme (object)) as for agent-subject verbs, but the role prototypicality of the arguments is reduced. In the sense of generalised thematic roles, while the experiencer-subject has properties of a Proto-Agent/Actor, it

is a less prototypical Proto-Agent/Actor than an agent-subject and the Stimulus/Theme object is a less prototypical Proto-Patient/Undergoer than a Patient object (Dowty 1991; Van Valin et al. 1997; Primus 1999)[1]. Further, while work on the Competition Model examined role assignment using an actor/subject identification probe (MacWhinney et al. 1984), much psycholinguistic literature has employed grammaticality or acceptability judgement tasks (e.g. Kolk et al. 2003; Hoeks et al. 2004, Bourguignon et al. 2012), rather than directly examining role assignment. While some research has used multiple question types within one experiment (e.g. Kyriaki et al. 2020), a comparison of actor versus undergoer identification probes (and their effects on comprehension) has not been directly examined to our knowledge.

With the present study, we aimed to address these gaps in our knowledge of cue use for sentence comprehension by comparing comprehension strategies in default configurations (sentences with transitive action verbs, agent-subject) with configurations differing mildly from the default (sentences with experiencer-subject verbs). A cross-linguistic perspective is taken by comparing English and German to consider potential effects of word order (in)flexibility. Further, we vary the comprehension task focus – probing for the actor or undergoer – to systematically study effects of the experimental environment on cue utilisation. We aimed to increase our understanding of the degree to which cue weightings shift during sentence comprehension based on both language-internal factors and the speaker's current goals (as determined by task focus). In the following, we first describe cross-linguistic differences in sentence interpretation, focusing on the languages examined in the present study: English and German. We then discuss previous examinations of role assignment in sentences with non-default verb classes and introduce thematic reversal anomalies (TRAs) as a paradigm for probing sentence interpretation strategies, before introducing the present study.

1.1. Cues to sentence comprehension across languages

Based on the comprehension of sentences involving transitive action verbs, MacWhinney et al. (1984) rank the cues of animacy and word order for thematic interpretation in English: Word Order > Animacy, and in German: Animacy > Word Order. The relative influence of these cues is exemplified in the results of psycholinguistic studies (for review, see Bornkessel-Schlesewsky et al. 2011). The strength of word order as a cue for English is shown where undergoer/object first sentences are less frequent (Hopper & Thompson 1980; Ferreira 1994), can take longer to process (King & Just 1991; Ferreira 2003), and lead to lower comprehension accuracy (King & Just 1991;

[1] Initial evidence from causative object-experiencer verbs in German suggests that word order preferences are preserved in this case (Scheepers et al. 2000). However, due to the causative nature of these verbs, their argument structure (Causer (subject) > Experiencer (object)) aligns more closely with that of action verbs (e.g. Grimshaw 1990). Additionally, the experiments reported by Scheepers and colleagues compared causative object-experiencer verbs with stative subject-experiencer verbs, thus contrasting word order preferences for two non-default verb classes.

Ferreira 2003). Animacy also contributes to English sentence comprehension, as speakers tend to place animates in the subject role and prefer agents or experiencers to take the subject rather than object role (Ferreira 1994).

In German, animacy cues speakers to a greater degree, with a preference for the animate participant in the actor role and the inanimate participant in the undergoer role (MacWhinney et al. 1984; Van Nice & Dietrich 2003; Chan et al. 2009). Word order also influences German sentence comprehension, with a preference for actor/subject-first constructions, shown by longer reading times (Fanselow et al. 2002), and decreased judgement accuracy (Bader & Meng 1999) for undergoer/object-first sentences. During sentence comprehension, morphosyntactic cues are quickly integrated and can overcome the word order preference for actor/subject-first constructions (Hopp 2006; Jackson 2008). Further, multiple permissible word orders (e.g. object-initial constructions) make this cue unreliable for interpretation (MacWhinney et al. 1984). Cue prominence may be modulated by the (mis)alignment of thematic, grammatical function, and/or case hierarchies (Schlesewsky & Bornkessel 2003; Bornkessel et al. 2004).

1.2. Thematic Reversal Anomalies to investigate role assignment

Thematic Reversal Anomalies (TRAs) are a linguistic manipulation that effectively probes sentence comprehension and role assignment. TRAs, where the two noun phrases (NPs) of a plausible transitive clause are reversed to elicit a semantic anomaly, allow for direct comparison between thematically plausible and implausible sentences (Kolk et al. 2003; Kuperberg et al. 2003; Hoeks, Stowe & Doedens 2004; Kim & Osterhout 2005; Van Herten et al. 2005). Thus, TRAs provide a strong test case for the ability of cues other than word order to "work against" the dominant order cue in English. We will capitalise on this property of TRAs in the present study. In addition, these manipulations have played a prominent role in the neurolinguistic literature, as the neurophysiological response to thematically plausible (1a) versus implausible (1b) sentences challenged previous assumptions about the neural language processing architecture (see Van de Meerendonk et al. 2009 for a review).

(1) Agent- subject	a.	The gardener will plough the soil with his tools.
	b.	The soil will *plough the gardener with his tools.
(2) Experiencer- subject	a.	Her boss will value her work for many years.
	b.	Her work will *value her boss for many years.
(3) Experiencer – object	a.	Her work will satisfy her boss for many years.
	b.	Her boss will satisfy her *work for many years.

(* asterisk marks point of violation; examples from Bourguignon et al. 2012).

Thematically implausible sentences elicit an animacy violation (marked *), which evokes a neural (event-related potential, ERP) response. Where experiencer-subject (2b) and experiencer-object (3b) violations are directly compared, similar ERPs are observed at the point of animacy violation: for experiencer-subject, at the verb, and for experiencer-object, at the second NP (Paczynski & Kuperberg 2011; Bourguignon et al. 2012). This neural response is qualitatively different compared to that elicited by violations of agent-subject verb constructions (1b) (Kolk et al. 2003; Kuperberg et al. 2003; Hoeks et al. 2004; Kim & Osterhout 2005; Van Herten et al. 2005; Bourguignon et al. 2012; Kyriaki et al. 2020). The observed ERP differences to violations of differing verb classes support the conception that agent and experiencer verbs are comprehended differentially at the neural level. Behavioural results also show higher grammaticality judgement and comprehension accuracy for agent-subject over experiencer-subject sentences (Bourguignon et al. 2012; Kyriaki et al. 2020). However, response accuracy results cannot provide insight into role assignment.

Cross-linguistic comparisons of TRA violations also align with the CM's proposed differences in relative cue weightings across languages (see Bornkessel-Schlesewsky et al. 2011 for overview). The neural response elicited by agent-subject verb (ASV) violations qualitatively differs between English and German (Bornkessel-Schlesewsky et al. 2011). In English, the prominent cue of word order leads to the implausible interpretation of an inanimate acting upon an animate (Kuperberg et al. 2003; Kim & Osterhout 2005). In German, TRA violations are resolvable in some cases, where flexible word order allows the possibility of reinterpretation (Bornkessel et al. 2002; Bornkessel-Schlesewsky et al. 2011). Here, where case marking is ambiguous, the cue strength of animacy in German can lead an ASV TRA to be resolved with reanalysis to an object-first interpretation (Bornkessel-Schlesewsky et al. 2011).

Prior research in English further shows a verb class difference in behavioural and neural responses to TRAs (Bourguignon et al. 2012; Kyriaki et al. 2020). While cross-linguistic comparisons for ASV TRAs have been examined across English and German (Bornkessel-Schlesewsky et al. 2011), other verb classes have not been compared. Non-default experiencer-object verbs have been shown to reduce the applicability of word order for comprehension in German (Schlesewsky & Bornkessel 2003; Bornkessel et al. 2004) but the effect of experiencer-subject verbs has not been directly investigated. The use of a less prototypical subject/Actor (experiencer) may lead to a stronger prominence of animacy (and weaker prominence of word order) as a cue, particularly in the comprehension of sentences with an inanimate NP1 and animate NP2. This hypothesis is based, in part, on the observation that experiencer-subjects tend to differ in their morphosyntactic properties from agent-subjects in prototypical transitive sentences (cf. Hopper & Thompson, 1980, for this notion), for example in terms of case marking patterns (see Bickel et al. 2014, for a review and empirical analysis). We thus suggest that, just as they tend to be associated with non-default morphosyntactic patterns, less prototypical subject/Actor arguments may lead to non-default comprehension strategies. Comparing the way

in which English and German speakers assign roles in agent-subject and experiencer-subject verb sentences would thus contribute towards the current understanding of sentence interpretation built upon neurolinguistic research. This also allows for a direct comparison of verb class across two languages which differentially rely on CM cues for interpretation (MacWhinney et al. 1984).

1.3. The present study

The present study is an internet-based sentence reading survey, aiming to investigate cross-linguistic differences in English and German thematic role assignment across agent- and experiencer-subject TRAs. While case marking is a cue to sentence comprehension it German, it is not always informative, such as in sentences where both NPs are feminine or plural (MacWhinney et al. 1984; Kempe & MacWhinney, 1999). German TRAs were designed with ambiguous case marking, allowing them to be resolvable with reconceptualization to an object-first word order. By constructing sentences with ambiguous case marking in this way, case marking can be rendered uninformative to interpretation while the sentence remains naturalistic. This also allows a more direct comparison of German to English (which does not use case marking). Example stimuli are shown in **Table 1**. We developed a comprehension task which prompts participants to identify either the actor or undergoer. This task was modelled after the task used in MacWhinney et al. (1984) who prompted participants for the actor/subject, the task described in Kyriaki et al. (2020) who prompted for either the actor or undergoer.

We predict (H1) that English speakers will display a predominantly word-order-based comprehension strategy, most often selecting the NP1 as the actor and NP2 as the undergoer regardless of animacy cues. We hypothesise (H2) that German speakers will display a predominantly animacy-based comprehension strategy, most often selecting the animate NP as the actor and inanimate NP as the undergoer regardless of word order cues. Lastly, we predict (H3) that object-initial interpretations in German will occur more frequently in sentences containing experiencers compared to sentences containing agents. We employed a $2 \times 2 \times 2 \times 2$ design with factors Verb Class (Agent-Subject Verb or Experiencer-Subject Verb), Animacy (of the NP1 and NP2; Animate-Inanimate or Inanimate-Animate), Question Type (Actor or Undergoer probe) and Language (English or German).

2. Methods

2.1. Participants

At the time of the study, all participants were currently residing in Australia, but German speakers had resided within a German-speaking country within the past two years (i.e., within the last two years, had moved from a German-speaking country to reside in Australia). To be eligible, participants were required to report no language disorder or intellectual impairment and had not learned any languages other than their native language (English or German) prior to the age of 5 years. Thirty-eight native English speakers (23 female, mean age 25.71 ± 0.87) took part in the English survey. Fourteen reported that they spoke a second language. Participants were asked to

Verb	Noun 1 – Noun 2 Animacy	Example Sentence
English		
ASV	A-I	The men will devour the meals in the restaurant.
ASV	I-A	The meals will *devour the men in the restaurant.
ESV	A-I	The children will enjoy the holidays in the village.
ESV	I-A	The holidays will *enjoy the children in the village.
German		
ASV	A-I	Die Männer werden die Mahlzeiten im the men.$_{NOM/ACC}$ will the meals.$_{NOM/ACC}$ in the Restaurant verschlingen. restaurant devour. 'The men will devour the meals in the restaurant.'
ASV	I-A	Die Mahlzeiten werden die Männer im the meals.$_{NOM/ACC.}$ will the men.$_{NOM/ACC}$ in the Restaurant verschlingen. restaurant devour.
ESV	A-I	Die Kinder werden. die Ferien im the children.$_{NOM/ACC}$ will the holiday.$_{NOM/ACC}$ in the Dorf genießen. village enjoy. 'The children will enjoy the holiday in the village.'
ESV	I-A	Die Ferien werden die Kinder im the holiday.$_{NOM/ACC}$ will the children.$_{NOM/ACC}$ in the Dorf genießen. village enjoy.

Table 1: Sentence conditions examined in the present study. Violations in English are marked by an asterisk (*). Translations of German I-A sentences are not included, as the interpretation of these sentences is the focus of the experimental manipulation.
ASV = agent-subject verb, ESV = experiencer-subject verb, A = animate, I = inanimate.

report their highest level of education. Two participants (5.26%) had not completed secondary school, 19 (50%) had completed secondary school, and 17 (44.74%) had completed a Bachelor's degree or higher. Twenty-six native German speakers (8 female, mean age 27.73 ± 1.02) took part in the German survey. All spoke English in addition to German, with four also speaking a third language. Thirteen (50%) had completed secondary school, and thirteen (50%) had completed a Bachelor's degree or higher.

2.2. Materials

2.2.1. Stimuli and comprehension task

English stimuli were drawn from Kyriaki et al. (2020) which were adapted from Bourguignon et al. (2012). See **Table 1** for example stimuli, with the full lists available open access (see data availability statement). Critical stimuli were constructed in pairs for Agent-Subject Verb (ASV) and Experiencer-Subject Verb (ESV) classes, with an animate NP as the first noun phrase and inanimate NP as the second noun phrase or vice versa. All English critical sentences followed the template *The NP will Verb the NP Prepositional Phrase/Adverbial Phrase*, while German critical sentences followed the template *The NP will the NP Prepositional Phrase/Adverbial phrase Verb*. Four stimulus lists were pseudorandomised, with a mirrored version of each list also created to rule out sequence effects. Each list contained 120 sentences (40 critical, 80 distractor) and comprehension questions. Lists were comprised of twenty ASV (10 A-I, 10 I-A) and twenty ESV (10 A-I, 10 I-A), forty distractor phrase structure violations (20 control, 20 violation, e.g. “He started to [thank his wife/*wife his thank]…”) and forty distractor semantic anomalies (20 control, 20 violation, e.g. “The bank will invest [the eggs/*the money]…”).

For the German survey, stimuli and survey instructions were translated by a native German speaker and checked by a second speaker. Some NPs were altered from the original English to similar feminine or plural NPs to allow for ambiguous case marking (e.g. “the labourer” altered to “Die Angestellte” [the employee.$_{F}$]). As it was not possible to translate phrase structure violation distractors from English accurately, German phrase structure violations were drawn from Frisch (2000). An important distinction between the English and German TRA violations is that a reinterpretation to a plausible object-first construction is possible for German, but not English, violation sentences. While the animacy cue may lead German speakers to a plausible interpretation, the strength of word order in English does not allow this possibility. In this sense, it is possible that German speakers can interpret the stimuli as plausible if animacy, and not word order, is used for interpretation.

The comprehension task consisted of three question types: probing for the Actor (e.g. Who/what will devour?/Wer/was wird verschlingen?), probing for the Undergoer (e.g. Who/what will be devoured?/Wer/was wird verschlungen werden?), or a yes/no probe (e.g. Was there a car?/Gab es ein Auto?). Below the question, two possible answers were presented: for Actor and Undergoer questions, two nouns (e.g. “meals” and “men”), and for the yes/no probe, “yes” and “no”. All critical sentences were followed by a pseudorandomly allocated Actor or Undergoer question, while distractor sentences were followed by an Actor, Undergoer, or yes/no probe question. The yes/no probe acted as a measure of general comprehension to compare across English and German.

2.2.2. Accounting for linguistic differences across items

As the semantic content of sentences differs across linguistic items, factors such as the pairwise cosine similarity (semantic relatedness) of nouns and verbs may influence the interpretation or

response to items. For each critical stimulus, pairwise cosine semantic similarities of nouns and verbs (e.g. of “men” and “devour”) were extracted from a semantic analysis space using package *LSAfun* v0.6.1 (Günther et al. 2015) in *R* v3.5.2 (R Core Team, 2018). The English (EN_100k_cbow) and German (dewak100k_cbow) semantic spaces each contained vectors for 100,000 words and were created using the cbow algorithm (as implemented in the word2vec model described in Mikolov et al. 2013). For further details on these semantic spaces and information on how to access them, see Günther et al. (2015).

The calculated pairwise cosine similarity of the first nouns and verbs was included in the statistical analysis as a covariate to account for differences across items (Sassenhagen & Alday 2016). This allows an estimation of the effect arising from the linguistic manipulation, while accounting for potential differences arising from semantic relatedness of nouns and verbs. Further, this practice is recommended as more statistically appropriate than null hypothesis significance testing of differences in linguistic features between stimuli/conditions (for discussion, see Sassenhagen & Alday 2016). Mean and standard deviation pairwise cosine similarities are listed in **Table 2**. While not analysed in this study, mean and standard deviations for orthographic length and word frequency are listed in **Table 3** for nouns and verbs of the critical sentences.

2.3. Procedure

Participants undertook the survey via an online survey provider (Qualtrics) in English or German. The first page presented the information sheet and consent form, and participants provided

Language	Noun 1 – Noun 2 Animacy	Verb Class	Mean (SD)
English	A-I	ASV	0.129 (0.127)
		ESV	0.100 (0.169)
	I-A	ASV	0.241 (0.166)
		ESV	0.108 (0.132)
German	A-I	ASV	0.086 (0.105)
		ESV	0.086 (0.118)
	I-A	ASV	0.250 (0.180)
		ESV	0.074 (0.127)

Table 2: Mean and standard deviation pairwise cosine similarity (semantic relatedness) between the first noun and verb of critical sentences.
A = animate, I = inanimate, ASV = Agent Subject Verb, ESV = Experiencer Subject Verb.

	English		German	
	Length (SD)	**Frequency (SD)**	**Length (SD)**	**Frequency (SD)**
Animate noun	7.11 (1.92)	11.57 (3.26)	8.26 (2.64)	10.48 (2.91)
Inanimate noun	6.81 (2.51)	11.71 (2.51)	8.69 (3.08)	11.49 (2.73)
ASV	5.85 (1.62)	11.59 (3.69)	9.00 (2.45)	11.74 (2.32)
ESV	5.75 (1.60)	12.83 (3.82)	8.19 (1.79)	12.53 (3.01)

Table 3: Mean and standard deviation orthographic length and frequency class across noun types (animate or inanimate) and verb types (ASV or ESV) for English and German.
English frequency calculated using British National Corpus v3, 2007 and German frequency calculated using Leipzig Corpora Collection, 2018. ASV = Agent Subject Verb, ESV = Experiencer Subject Verb.

consent by selecting "I agree". After providing demographic information to confirm eligibility, instructions were presented. Two example sentences and practise questions (one Actor, one Undergoer probe) were presented. For the experimental part of the survey, sentences and questions were presented individually page-by-page, with no option to return to a previous page. Participants were instructed to read each sentence once, before pressing the "next" button to proceed. The following page displayed a question about the previous sentence with two possible answers. Participants were instructed to select the answer they believed was most correct, then press "next". The survey lasted for approximately 20 minutes. At the survey's conclusion, participants were offered the opportunity to provide their details to receive an honorarium.

3. Data analysis

R version 3.5.2 (R Core Team, 2018) was used for statistical analyses with packages *tidyverse* v1.2.1 (Wickham, 2017), *car* v3.0.2 (Fox & Weisberg, 2011), *lme4* v1.1.21 (Bates et al. 2015) *effects* v4.1.0 (Fox, 2003; Fox & Weisberg, 2019), and *emmeans* v1.5.1 (Lenth, 2021). Plots were created in R using *ggplot2* v3.3.0 (Wickham et al. 2020) and *ggpubr* v0.3.0 (Kassambara, 2020), with *lmerOut* v0.5 (Alday, 2018) used to produce model output tables.

Behavioural responses to the critical stimuli (which noun was selected) were analysed using a generalized linear mixed effect model (GLMM) fit by maximal likelihood. Fixed effects included Noun Animacy (A-I or I-A), Verb Class (ASV or ESV), Question Type (Actor or Undergoer probe) and Language (English or German) and their interactions. Pairwise cosine similarity values (between the first noun and verb) were included as a covariate (main effect). Intercepts were grouped by Participant ID and Item, and random slopes for the effect of Noun Animacy were constructed by Participant ID and Item. Behavioural response was specified as the dependent variable.

Categorical variables were encoded with sum contrasts (ANOVA-style encoding) where the intercept of the model is the grand mean (for further information, see Schad et al. 2020 and Appendix A). Main effects and interaction terms were assessed using Wald tests from Type-II ANOVA tables obtained using *car* (Fox & Weisberg, 2011). *P*-values for model terms were derived using Satterthwaite's method for approximating degrees of freedom. Post-hoc contrasts (using the Tukey correction for multiple comparisons) were obtained using *emmeans* (Lenth, 2020). Where models are plotted, error bars are 83% confidence intervals, where non-overlapping intervals indicate a significant difference at $\alpha = .05$ (MacGregor-Fors & Payton, 2013). We also conducted a secondary analysis on responses to the yes/no probes for distractor stimuli to compare general comprehension across languages (see Appendix B).

4. Results

Descriptive statistics for response choices across languages and verb classes are summarised in **Tables 4** and **5** for actor and undergoer probes, respectively.

Verb Class	Noun Animacy	English Mean (SD)	German Mean (SD)
ASV	A-I	86.54% (34.21%)	71.74% (45.19%)
	I-A	70.30% (45.81%)	53.62% (50.05%)
ESV	A-I	78.71% (41.04%)	68.38% (46.67%)
	I-A	62.89% (48.44%)	34.33% (47.66%)

Table 4: Actor probe: Mean and standard deviation (SD) selection of the NP1 in response to the actor probe for English and German, by verb class and noun animacy.
A = animate, I = inanimate, ASV = Agent Subject Verb, ESV = Experiencer Subject Verb.

Verb Class	Noun Animacy	English Mean (SD)	German Mean (SD)
ASV	A-I	80.81% (39.49%)	65.57% (47.71%)
	I-A	60.11% (49.10%)	36.89% (48.45%)
ESV	A-I	78.65% (41.09%)	50.81% (50.20%)
	I-A	57.53% (49.56%)	38.10% (48.76%)

Table 5: Undergoer probe: Mean and standard deviation (SD) selection of the NP2 in response to the undergoer probe for English and German, by verb class and noun animacy.
A = animate, I = inanimate, ASV = Agent Subject Verb, ESV = Experiencer Subject Verb.

Statistical analysis (see **Table 6**) revealed a significant four-way Noun Animacy × Verb Class × Question Type × Language interaction effect along with a three-way interaction effect of Noun Animacy × Question Type × Language on response choice. All model parameters also showed significant main effects, as reported in **Table 6**, with the model summary included in Appendix A. There was a main effect of Verb Class, such that response choice followed a more word-order-based strategy for ASV compared to ESV sentences. There was also a main effect of Noun Animacy where

Parameters	**χ2**	**df**	***p***
Verb Class	19.89	1	<0.001*
Noun Animacy	28.40	1	<0.001*
Question Type	10.50	1	0.001*
Language	21.27	1	<.001*
Pairwise Cosine Similarity	4.67	1	0.031*
Verb Class × Noun Animacy	0.00	1	0.945
Verb Class × Question Type	1.10	1	0.295
Noun Animacy × Question Type	0.05	1	0.815
Verb Class × Language	0.28	1	0.600
Noun Animacy × Language	1.06	1	0.303
Question Type × Language	0.38	1	0.539
Verb Class × Noun Animacy × Question Type	2.08	1	0.149
Verb Class × Noun Animacy × Language	0.21	1	0.648
Verb Class × Question Type × Language	0.56	1	0.454
Noun Animacy × Question Type × Language	5.13	1	0.023*
Verb Class × Noun Animacy × Question Type × Language	8.09	1	0.004*

Table 6: Type II Wald chi square (χ2) tests for main effects and interaction terms examining the relationship between Verb Class (agent subject verb, experiencer subject verb), Noun Animacy (animate-inanimate, inanimate-animate), Question Type (actor probe, undergoer probe) and Language (English, German) on the response choice to comprehension questions.

Note. Asterisk (*) denotes statistical significance at $p < 0.05$.

Animate-first sentences had a higher word-order response strategy than Inanimate-first sentences. The main effect of Question Type showed that Actor probes had a higher word-order strategy than Undergoer probes. Lastly, there was a main effect of Language, where English-speakers had a stronger word-order strategy than German speakers. **Figures 1** and **2** present graphical representations of the modelled results. The secondary analysis of yes/no probes to distractor sentences found no significant main or interaction effects including Language, indicating similar general comprehension across languages (see Appendix B).

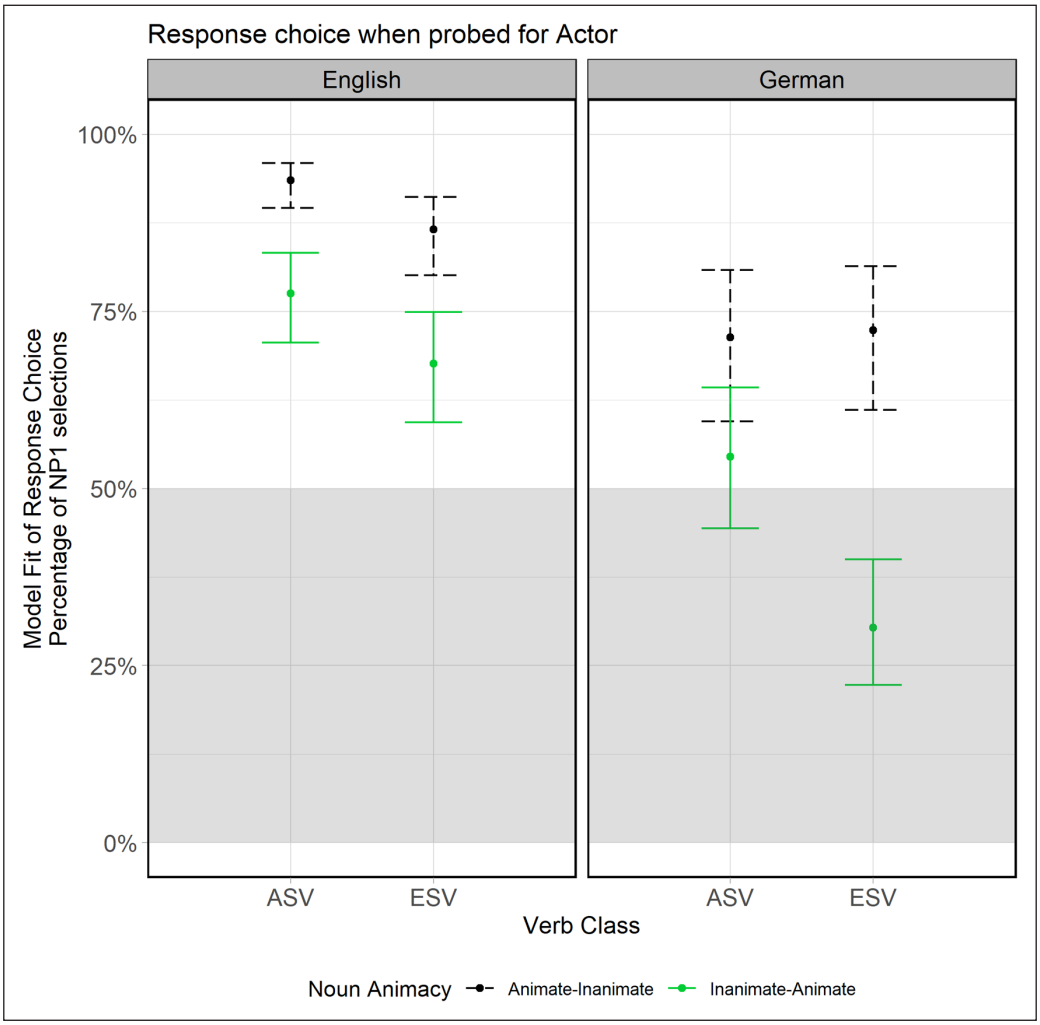

Figure 1: Modelled interaction of Noun Animacy, Verb Class, and Language on response choice when probed for the Actor. Error bars are 83% confidence intervals around model fit, where non-overlapping intervals indicate a significant difference at $\alpha = .05$. The grey shaded area denotes a percentage of NP1 selection below 50%, indicating 50% or more NP2 choices when probed for the Actor and a lesser prominence of word order as a cue to thematic role assignment. A matched plot of raw data is available in Appendix C. ASV = Agent-subject verb, ESV = Experiencer-subject verb.

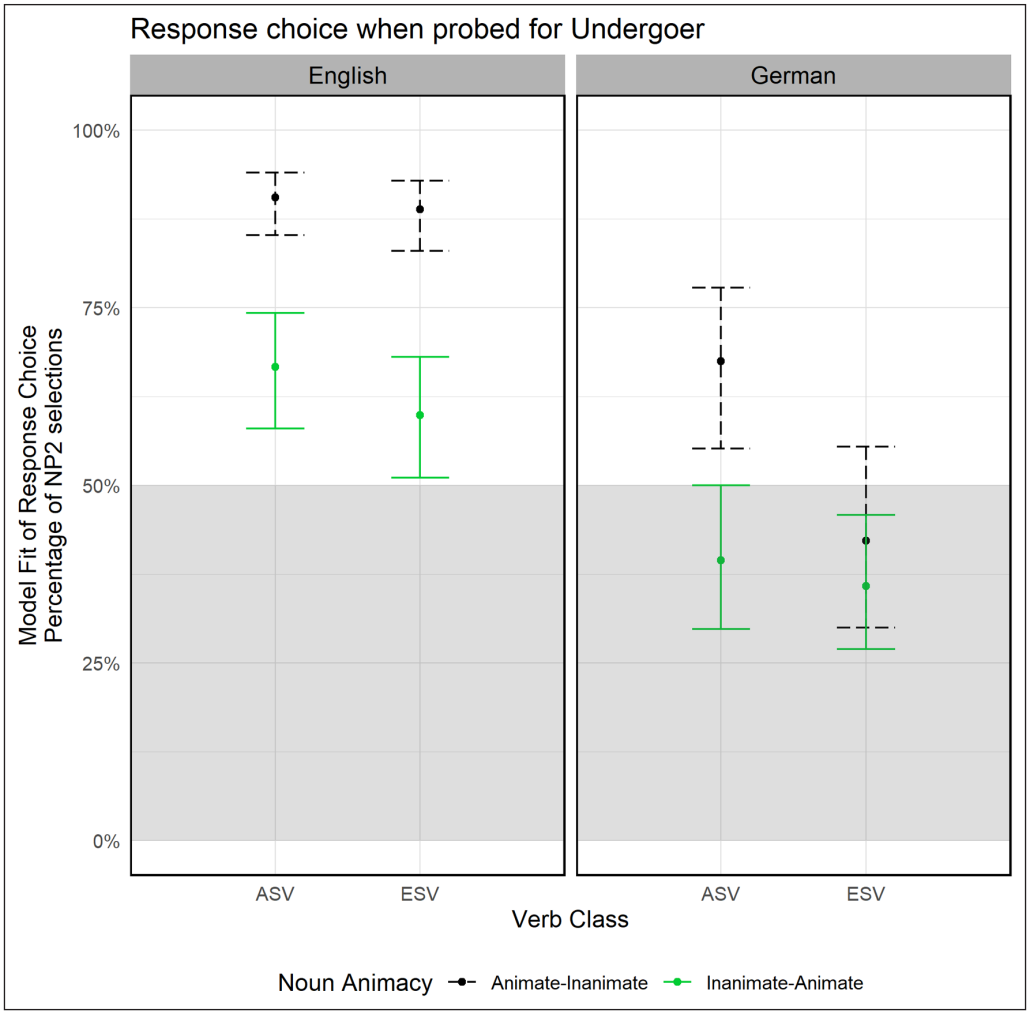

Figure 2: Modelled interaction of Noun Animacy, Verb Class, and Language on response choice when probed for the Undergoer. Error bars are 83% confidence intervals around model fit, where non-overlapping intervals indicate a significant difference at $\alpha = .05$. The grey shaded area denotes a percentage of NP2 responses below 50%, indicating 50% or more NP1 choices when probed for the Undergoer, and a lesser prominence of word order as a cue to thematic role assignment. A matched plot of raw data is available in Appendix C. ASV = Agent-subject verb, ESV = Experiencer-subject verb.

In view of the overarching interactions involving language, we discuss the results for English and German in turn in the following. Based on the significant four-way interaction of Language × Verb Class × Noun Animacy × Question Type, the post-hoc pairwise comparisons discussed below were calculated for levels of Noun Animacy (A-I, I-A), by each level of Verb Class (ASV, ESV), Question Type (actor probe, undergoer probe), and Language (English, German).

4.1. English response choice

The Noun Animacy × Verb Class × Question Type × Language interaction is visualised in **Figures 1** and **2** for actor and undergoer probes, respectively. For English speakers, post-hoc contrasts indicate that A-I sentences elicited significantly higher word-order-based comprehension compared to I-A sentences regardless of question type and verb class (for the actor probe, ASV $p < 0.001$, ESV $p = 0.003$; for the undergoer probe, ASV $p < 0.001$, ESV $p < 0.001$).

4.2. German response choice

For German, the pattern of results differed more widely across verb class and question type. For actor probes, ASV sentences showed no significant difference ($p = 0.08$) in response between A-I and I-A sentences, indicating a similar response pattern regardless of noun animacy. For ESV, A-I constructions led to a significantly higher proportion of word-order-based interpretations, while I-A led to more animacy-based interpretations ($p < 0.001$). When probed for the undergoer, ASV A-I constructions elicited significantly more word-order-based responses than did I-A constructions ($p = 0.007$), but there was no difference between noun animacy conditions for ESV sentences ($p = 0.53$). Here, both A-I and I-A ESV constructions elicited a response strategy where the NP1 was more often selected as the undergoer (regardless of animacy).

5. Discussion

We examined differences in role assignment for English and German, using thematic reversal anomalies (TRAs) to investigate the effects of noun animacy and verb class. We also explored the retrospective influence of task focus on role assignment, where in some cases, the focus of the comprehension question led to reinterpretation of sentences. While cross-linguistic (e.g. Kim & Osterhout 2005; Van de Meerendonk et al. 2009, Bornkessel-Schlesewsky et al. 2011) and verb class (Bourguignon et al. 2012, Kyriaki et al. 2020) differences in the neurophysiological response to TRAs have been observed, role assignment has not been examined in detail. We selected English and German for investigation as both cue prominence (MacWhinney et al. 1984) and the observed neurophysiological response to TRAs (Bornkessel-Schlesewsky et al. 2011) contrast between the two languages. As predicted (H1), English speakers displayed a predominantly word-order-based comprehension strategy. In comparison, German speakers showed a higher proportion of responses in line with an animacy-based strategy. However, responses were not predominantly animacy-based in all conditions as hypothesised (H2), with ASV I-A constructions (*"The pyramids will visit…"*) eliciting a word-order-based strategy in more than half of responses. This unexpected prominence of word order for German speakers could potentially be due to their surrounding environment, as they lived in an English-speaking country. Lastly, as predicted (H3), animacy-driven object-initial interpretations in German were more likely in ESV sentences

compared to ASV. In both languages, responses to critical questions were not at ceiling level of performance (e.g. identification of the NP1 as the actor in an ASV A-I sentence, where all cues align, occurred in ~87% of English and ~72% of German responses). However, this was not surprising given response patterns to similar stimuli and questions in previous research where even simple grammatical judgements were not at ceiling performance (Bourguignon et al. 2012; Kyriaki et al. 2020). General comprehension (as measured by the yes/no questions) was high and did not significantly differ as a function of language (see Appendix B).

For both languages, noun animacy, verb class, and question type interacted to influence participants' response patterns relating to the cues of word order and animacy (see **Figure 3** for schematic). The overall pattern of results shows that the alignment of certain conditions appears to affect the prominence of word order versus animacy for interpretation. Where noun animacy was A-I and thus in line with an order-based strategy, interpretations predominantly followed the word order cue, compared to I-A constructions where word order may lead to an implausible interpretation, thus promoting the prominence of animacy. A similar pattern was observed for verb class, where default agent-subject verbs promote word order as a cue and experiencer-subjects increase the prominence of animacy. The focus of the question – probing for the actor or undergoer – also showed an effect on role assignment. In general, the probed role was more

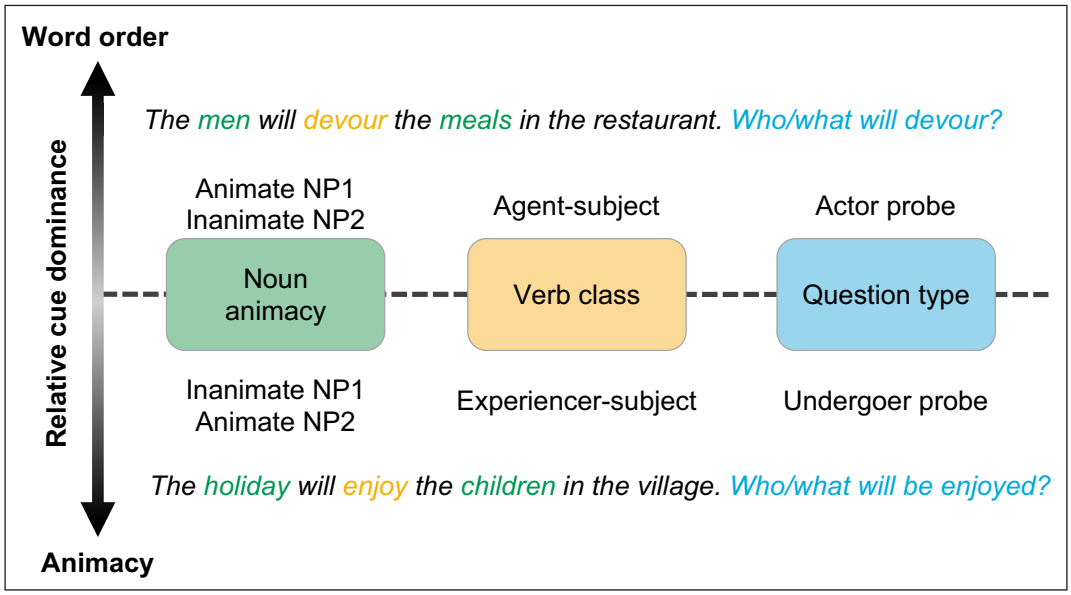

Figure 3: Schematic of the relative cue strength for word order and animacy, and the influence of noun animacy, verb class, and question type. For each factor, one level was associated with a higher word order cue dominance (above dashed line) and one level was associated with a strengthening of the animacy cue (below dashed line). An example sentence and question are presented that represent a sentence context where word order is a more prominent cue (above dashed line) or where the prominence of animacy is increased (below dashed line).

likely to be equated with the first argument, with this effect stronger in German. The influence of each experimental factor appears to be "cumulative" in promoting the relative prominence of a cue for thematic interpretation. Results will be discussed in the context of the Competition Model, before the role of question type and verb class are examined in more detail. We will then discuss the contribution of our findings to current theoretical models.

5.1. Competition Model cue prominence

In line with Competition Model predictions (CM; Bates et al. 1982; MacWhinney et al. 1984), English and German speakers relied on both word order and animacy as cues to interpretation, but to different degrees. The dominant cue to English role assignment was word order for both ASV and ESV sentences. English speakers showed a strong tendency to select the NP1 as the actor, matching the results of Bates et al. (1982), who found that in noun-verb-noun constructions, English speakers select the NP1 as the actor most often. We also observed a slight animacy preference corresponding with Dowty's (1991) thematic proto-roles. English speakers preferred animate NPs as the actor/proto-agent (and inanimate NPs as the undergoer/proto-patient). When probed for the actor, animate NP1s were selected more often than inanimate NP1s. This illustrates that subordinate cues influence thematic interpretation concurrently to dominant cues even under circumstances where the dominant cue is thought to shape interpretation in a quasi-deterministic manner.

For German speakers, the ambiguous case marking of NPs (due to feminine gender or plural number) rendered case uninformative to role assignment. Both animacy and word order cued sentence interpretation, with animacy a prominent cue, but not completely dominant over word order as hypothesised. Nevertheless, animacy strongly determined interpretation: compared to English, inanimate NPs were selected as actors less often, and animate NPs were selected as undergoers less often. In addition, sentential and experimental context – here, question type and verb class – strongly influenced the relative prominence of both word order and animacy as cues to interpretation, which will be discussed below.

Overall, our findings support the CM's claim that there is no principled difference between cue weightings that represent tendencies and those that represent deterministic rules (MacWhinney et al. 1984, p. 129). However, they also go somewhat beyond the original formulation of the CM, in suggesting that cue weightings might be upgraded or downgraded in the context of other cues. As we discuss further below, the weighting of an animacy-based strategy over a word-order-based strategy was increased for sentences with non-default (experiencer-subject) verbs and with undergoer-focused questions, particularly where word order was more flexible (i.e. in German as opposed to English). This suggests a possible contextualisation of cue utilisation in a manner not directly envisaged by the CM. We will return to this point in the discussion of theoretical implications in section 5.4.

5.2. Question context retrospectively influences topicality

The focus of the comprehension question significantly affected response choice, and this was particularly apparent in German, where question type interacted with verb class and animacy. The nature of the comprehension question retrospectively influenced semantic interpretation, likely by promoting the NP1 to sentence topic (what the sentence "is about"; Reinhart 1981). In German, the topical entity is preferred in the sentence initial position (Rosengren 1993; Büring 1999; Frey 2005; Schumacher & Hung 2012). Accordingly, when probed for the undergoer of an I-A construction, animacy-driven object-initial interpretations were more common, with the inanimate NP1 reconceptualised as the undergoer and topic. Strikingly, a significant difference between verb classes was observed when noun animacy and question type did not align. Responses to ASV versus ESV significantly differed where the question probed for the actor of an I-A construction and where the question probed for the undergoer of an A-I construction, with animacy a stronger cue in both animacy conditions for ESV. These verb class effects will be discussed in the following section.

In contrast to German, we did not see a similar topicality effect in English, which we posit to be due to the cue prominence (and inflexibility) of word order in English. Our findings are supported by language-specific differences in language production as a function of word order flexibility (for review, see Myachykov et al. 2011). Here, when the experimental context draws attentional focus to the undergoer of a transitive event, English speakers are more likely to produce a passive sentence to describe the event. Languages with a flexible word order (e.g. Russian), by contrast, show a higher number of object-first sentences and almost no passives in similar manipulations (Myachykov et al. 2011). Our results show that similarly to language production, language comprehension is influenced by the attentional affordance of the experimental context (here, question focus) differentially across languages. These results are also in line with previous assertions that experimental context can influence responses to linguistic violations (Osterhout & McKinnon 1996; Osterhout et al. 2002; Sassenhagen et al. 2014; Kyriaki et al. 2020).

5.3. Non-default verbs promote subordinate cues

Verb class effects on role assignment were highly influenced by the interaction of noun animacy and question focus for German. The general pattern of results was similar for ASV sentences across noun animacy conditions and question types. However, in German ESV sentences, the non-default verb class was associated with more frequent reinterpretation to object-first constructions, particularly when the question probed for the undergoer. English speakers showed no significant differences in response across verb classes, although the percentage of word-order-based responses was higher for ASV over ESV sentences. This held across noun animacy (A-I and I-A) and question conditions (Actor and Undergoer probes). The reduced cue strength of word

order for ESV sentences suggests that the non-default verb class increased the prominence of the subordinate cue of animacy. These findings are supported by behavioural results of a recent neurophysiological study of English speakers (Kyriaki et al. 2020), which found that response accuracy was lower for ESV compared to ASV, indicating that participants processed the two verb classes differently.

Overall, we assume that the modulation of interpretations by verb class was likely driven by the divergence of experiencers from agents (for review, see Rozwadowska 2017). While agent- and experiencer-subjects are both required to be animate, the latter is a less prototypical actor/proto-agent (Dowty 1991; Van Valin et al. 1997; Primus 1999) and this appears to promote subordinate cues for interpretation.

5.4. Theoretical implications

Our results show that, within the broader interpretation patterns shaped by dominant cues (e.g. word order in English), subordinate cues (e.g. animacy in English) also have a measurable effect on sentence interpretation. This is in line with the assumption of grammatical frameworks that allow for probabilistic constraint rankings (e.g. stochastic Optimality Theory; Boersma & Hayes 2001) or hard and soft constraints (e.g. Bresnan et al. 2001; Sorace & Keller 2005). Intriguingly, the relative weighting of subordinate cues was enhanced in the presence of a non-default verb class (experiencer-subject verbs) even though the thematic role hierarchy between subject and object was the same as in default (action) verbs. This indicates that – in addition to their well-known morphosyntactic differences from default verb classes (e.g. Primus, 1999) – non-default verb classes also engender non-default comprehension strategies, and this applies regardless of the "degree of deviance" from the default. While promotion of non-default strategies was stronger in German in the present study, likely due to the availability of flexible word order, it was also observable in English.

Similar considerations appear to hold for "non-default questions", i.e. those targeting the Undergoer rather than the Actor role. Strikingly, although the type of question encountered was not clear until after sentence presentation, Undergoer questions appeared to boost the prominence of the Undergoer (likely by enhancing topicality, as discussed above) and again promote the use of non-dominant interpretation strategies. This effect parallels that of non-default verb class in that it was stronger where flexible word orders were available (German) as opposed to when they were not (English). Moreover, verb class and question type coalesced in German, leading to the highest proportion of interpretations differing from the default.

Taken together, these observations suggest that cue weightings are surprisingly flexible and can be shifted by a number of contextual and environmental influences. These can be language internal, such as the use of non-default verb classes, or task-related, such as the use of different

question types. In all cases, deviations from default configurations seem to promote the use of non-default interpretation strategies. Further, the present results suggest a complex interaction of cues indicative of a weighting of constraints that goes beyond a strictly hierarchical (ordinal) ranking and is more in line with probabilistic models of grammar and language processing (cf. Jurafsky, 2003).

6. Conclusion

The present study shows that non-default experiencer-subjects elicit differential role assignment strategies to default agent-subjects, in both English and German. Further, experimental context – in this case, the task focus– significantly influences sentence comprehension. Future research should continue to investigate the differences in agent- and experiencer-subject verb role assignment across languages. The retrospective influence of task on comprehension should also be considered when designing future paradigms.

Abbreviations

A-I	Animate-Inanimate
ASV	Agent Subject Verb
CM	Competition Model
EOV	Experiencer Object Verb
ESV	Experiencer Subject Verb
GLMM	Generalised Linear Mixed (Effect) Model
I-A	Inanimate-Animate
NP	Noun Phrase
NP1	Noun Phrase 1
NP2	Noun Phrase 2
TRA	Thematic Reversal Anomaly

Data accessibility statement

Stimuli and datasets are in a publicly accessible repository: The datasets generated for this study and the analysis code can be found in an Open Science Framework repository. Link: https://osf.io/pu5kg/

Additional files

The additional files for this article can be found as follows:

- **Appendix A.** Model summary for the mixed effects model. DOI: https://doi.org/10.16995/glossa.5728.s1
- **Appendix B.** Response accuracy to distractor sentences. DOI: https://doi.org/10.16995/glossa.5728.s2
- **Appendix C.** Raw response data to critical sentences. DOI: https://doi.org/10.16995/glossa.5728.s3

Ethics and consent

Participants provided informed consent and were offered an honorarium for their participation. The study was approved by the University of South Australia Human Ethics Committee (approval no. 200107).

Acknowledgements

We would like to thank Emma Harrell for assisting in data pre-processing, and Andrew Corcoran, Amanda Santamaria, and Zachariah Cross for helpful discussions. We thank three anonymous reviewers for their helpful feedback on an earlier version of the manuscript.

Funding information

L.K. acknowledges the support of an Australian Government Research Training Program Scholarship (212791). I.B-S. acknowledges the support of an Australian Research Council Future Fellowship (FT160100437).

Competing interests

The authors have no competing interests to declare.

Author contributions

L.K., M.S., and I.B-S designed the research. L.K. conducted the research. L.K. analysed the data. L.K., M.S., and I.B-S wrote the paper.

References

Alday, Phillip. M. 2018. lmerOut: LaTeX Output for Mixed Effects Models with lme4. R package version 0.5. https://bitbucket.org/palday/lmerout.

Bader, Markus & Meng, Michael. 1999. Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research* 28(2). 121–143. DOI: https://doi.org/10.1023/A:1023206208142

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Bates, Elizabeth & MacWhinney, Brian. 1989. *Functionalism and the competition model*. In Bates, Elizabeth & MacWhinney, Brian (eds.), *The Crosslinguistic Study of Sentence Processing*, 3–73. Cambridge: Cambridge University Press.

Bates, Elizabeth & McNew, Sandra & MacWhinney, Brian & Devescovi, Antonella & Smith, Stan. 1982. Functional constraints on sentence processing: A cross-linguistic study. *Cognition* 11(3). 245–299. DOI: https://doi.org/10.1016/0010-0277(82)90017-8

Belletti, Adriana & Rizzi, Luigi. 1988. Psych-verbs and þeta-theory. *Natural Language & Linguistic Theory*, 291–352. DOI: https://doi.org/10.1007/BF00133902

Bickel, Balthasar & Zakharko, Taras & Bierkandt, Lennart, & Witzlack-Makarevich, Alena. 2014. Semantic role clustering: An empirical assessment of semantic role types in non-default case assignment. *Studies in Language* 38(3). 485–511. DOI: https://doi.org/10.1075/sl.38.3.03bic

Boersma, Paul & Hayes, Bruce. 2001. Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32(1). 45–86. DOI: https://doi.org/10.1162/002438901554586

Bornkessel, Ina & McElree, Brian & Schlesewsky, Matthias & Friederici, Angela D. 2004. Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language* 51(4). 495–522. DOI: https://doi.org/10.1016/j.jml.2004.06.011

Bornkessel, Ina, & Schlesewsky, Matthias & Friederici, Angela D. 2002. Grammar overrides frequency: Evidence from the online processing of flexible word order. *Cognition* 85(2). B21–B30. DOI: https://doi.org/10.1016/S0010-0277(02)00076-8

Bornkessel-Schlesewsky, Ina & Kretzschmar, Franziska & Tune, Sarah & Wang, Luming & Genç, Safiye & Philipp, Markus, & Roehm, Dietmar & Schlesewsky, Matthias. 2011. Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language* 117(3). 133–152. DOI: https://doi.org/10.1016/j.bandl.2010.09.010

Bornkessel-Schlesewsky, Ina & Schlesewsky, Matthias. 2020. *Cross-linguistic neuroscience of language*. In Poeppel, David & Mangun, George R. & Gazzaniga, Michael S. (eds.), *The Cognitive Neurosciences*, 6th ed. Cambridge: MIT Press.

Bourguignon, Nicolas & Drury, John E. & Valois, Daniel & Steinhauer, Karsten. 2012. Decomposing animacy reversals between agents and experiencers: An ERP study. *Brain and Language* 122(3). 179–189. DOI: https://doi.org/10.1016/j.bandl.2012.05.001

Bresnan, Joan & Dingare, Shipra & Manning, Christopher D. 2001. *Soft constraints mirror hard constraints: Voice and person in English and Lummi*. In Butt, Miriam & King, Tracy H. (eds.), *Proceedings of the LFG '01 Conference*. Stanford: CSLI Publications.

Büring, Daniel. 1999. *Topic*. In Bosch, Peter & van der Sandt, Rob (eds.), *Focus: Linguistic, cognitive, and computational perspectives*, 142–165. Cambridge: Cambridge University Press.

Chan, Angel & Lieven, Elena & Tomasello, Michael. 2009. Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics* 20(2). 267–300. DOI: https://doi.org/10.1515/COGL.2009.015

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619. DOI: https://doi.org/10.1353/lan.1991.0021

Fanselow, Gisbert & Kliegl, Reinhold & Schlesewsky, Matthias. 2002. *Processing Difficulty and Principles of Grammar*. In Kemper, Susan & Kliegl, Reinhold (eds.), *Constraints on Language: Aging, Grammar, and Memory*, 170–200. Boston: Kluwer Academic Publishers. DOI: https://doi.org/10.1007/0-306-46902-2_7

Ferreira, Fernanda. 1994. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language* 33(6). 715–736. DOI: https://doi.org/10.1006/jmla.1994.1034

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

Fox, John. 2003. Effect Displays in R for Generalised Linear Models. *Journal of Statistical Software* 8(15). 1–27. http://www.jstatsoft.org/v08/i15/. DOI: https://doi.org/10.18637/jss.v008.i15

Fox, John & Weisberg, Sanford. 2019. An {R} Companion to Applied Regression, Third Edition. Thousand Oaks: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Frey, Werner. 2005. *Zur Syntax der linken Peripherie im Deutschen*. In d'Avis, Franz-Josef (ed.), *Deutsche Syntax: Empirie und Theorie* 147. 147–171. Göteborg: Acta Universitatis Gothoburgensis.

Frisch, Stefan. 2000. *Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen*. Max Planck Institute of Cognitive Neuroscience Leipzig dissertation.

Gattei, Carolina A. & Dickey, Michael W. & Wainselboim, Alejandro J. & París, Luis. 2015. The thematic hierarchy in sentence comprehension: A study on the interaction between verb class and word order in Spanish. *Quarterly Journal of Experimental Psychology* 68(10). 1981–2007. DOI: https://doi.org/10.1080/17470218.2014.1000345

Gattei, Carolina A. & Sevilla, Yamila & Tabullo, Ángel J. & Wainselboim, Alejandro J. & París, Luis. & Shalom, Diego E. 2018. Prominence in Spanish sentence comprehension: an eye-tracking study. *Language, Cognition and Neuroscience* 33(5). 587–607. DOI: https://doi.org/10.1080/23273798.2017.1397278

Grimshaw, Jane. 1990. *Argument structure*. The MIT Press.

Günther, Fritz & Dudschig, Carolin & Kaup, Barbara. 2015. LSAfun – An R package for computations based on Latent Semantic Analysis. *Behavior Research Methods* 47(4). 930–944. DOI: https://doi.org/10.3758/s13428-014-0529-0

Hoeks, John C. J. & Stowe, Laurie A. & Doedens, Gina. 2004. Seeing words in context: the interaction of lexical and sentence level information during reading. *Cognitive Brain Research* 19(1). 59–73. DOI: https://doi.org/10.1016/j.cogbrainres.2003.10.022

Hopp, Holger. 2006. Syntactic features and reanalysis in near-native processing. *Second Language Research* 22(3). 369–397. DOI: https://doi.org/10.1191/0267658306sr272oa

Hopper, Paul J. & Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language*, 251–299. DOI: https://doi.org/10.1353/lan.1980.0017

Jackson, Carrie. 2008. Proficiency level and the interaction of lexical and morphosyntactic information during L2 sentence processing. *Language Learning* 58(4). 875–909. DOI: https://doi.org/10.1111/j.1467-9922.2008.00481.x

Jurafsky, D. 2003. *Probabilistic modeling in psycholinguistics: Linguistic comprehension and production*. In Bod, Rens & Hay, Jennifer & Jannedy, Stefanie (eds.), *Probabilistic Linguistics*. Cambridge: MIT Press.

Kassambara, Alboukadel. 2020. ggpubr: "ggplot2" Based Publication Ready Plots (Version 0.3.0) [Computer software]. https://CRAN.R-project.org/package=ggpubr.

Kempe, Vera & MacWhinney, Brian. 1999. Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes* 14. 129–171. DOI: https://doi.org/10.1080/016909699386329

Kim, Albert & Osterhout, Lee. 2005. The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language* 52(2). 205–225. DOI: https://doi.org/10.1016/j.jml.2004.10.002

King, Jonathan & Just, Marcel A. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30(5). 580–602. DOI: https://doi.org/10.1016/0749-596X(91)90027-H

Kolk, Herman H. J. & Chwilla, Dorothee J. & Van Herten, Marieke & Oor, Patrick J. W. 2003. Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language* 85(1). 1–36. DOI: https://doi.org/10.1016/S0093-934X(02)00548-5

Kuperberg, Gina R. & Sitnikova, Tatiana & Caplan, David & Holcomb, Phillip J. 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research* 17(1). 117–129. DOI: https://doi.org/10.1016/S0926-6410(03)00086-7

Kyriaki, Louise & Schlesewsky, Matthias & Bornkessel-Schlesewsky, Ina. 2020. Semantic reversal anomalies under the microscope: Task and modality influences on language-associated event-related potentials. *European Journal of Neuroscience* 52(7). 3803–3827. DOI: https://doi.org/10.1111/ejn.14862

Lenth, Russell V. 2021. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package, version 1.6.0. https://CRAN.R-project.org/package = emmeans.

MacGregor-Fors, Ian, & Payton, Mark E. 2013. Contrasting diversity values: statistical inferences based on overlapping confidence intervals. *PLoS One* 8(2). e56794. DOI: https://doi.org/10.1371/journal.pone.0056794

MacWhinney, Brian & Bates, Elizabeth & Kliegl, Reinhold. 1984. Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior* 23(2). 127–150. DOI: https://doi.org/10.1016/S0022-5371(84)90093-8

Mikolov, Tomas & Chen, Kai & Corrado, Greg, & Dean, Jeffrey. 2013. Efficient estimation of word representations in vector space. arXiv preprint. DOI: https://arxiv.org/abs/1301.3781v3

Myachykov, Andriy & Thompson, Dominic & Scheepers, Christoph & Garrod, Simon. 2011. Visual attention and structural choice in sentence production across languages. *Language and Linguistics Compass* 5(2). 95–107. DOI: https://doi.org/10.1111/j.1749-818X.2010.00265.x

Osterhout, Lee & McKinnon, Richard. 1996. On the language specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience* 8(6). 507. DOI: https://doi.org/10.1162/jocn.1996.8.6.507

Osterhout, Lee & Allen, Mark D. & Mclaughlin, Judith & Inoue, Kayo. 2002. Brain potentials elicited by prose-embedded linguistic anomalies. *Memory and Cognition* 30(8). 1304–1312. DOI: https://doi.org/10.3758/BF03213412

Paczynski, Martin & Kuperberg, Gina R. 2011. Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb-argument processing. *Language and Cognitive Processes* 26(9). 1402–1456. DOI: https://doi.org/10.1080/01690965.2011.580143

Primus, Beatrice. 1999. *Cases and Thematic Roles: Ergative, Accusative and Active*. Tiubingen: Max Niemeyer Verlag. DOI: https://doi.org/10.1515/9783110912463

Primus, Beatrice & Lindner, Katrin. 1994. Variation in grammar and first language acquisition: A new concept of parameter in universal grammar. *How Tolerant Is Universal Grammar? Essays on Language Learnability and Language Variation*, 185–204. DOI: https://doi.org/10.1515/9783111634777.185

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy I. *Philosophica Anc Studia Philosophica Gandensia Gent* 27(1). 53–94.

Rosengren, Inger. 1993. *Wahlfreiheit mit Konsequenzen–Scrambling, topikalisierung und FHG im dienste der informationsstrukturierung*. In Reis, Marga (ed.), *Wortstellung und informationsstruktur*, 251–312. Tübingen: Niemeyer. DOI: https://doi.org/10.1515/9783111658469.251

Rozwadowska, Bożena. 2017. Psychological verbs and psychological adjectives. *The Wiley Blackwell Companion to Syntax, Second Edition*, 1–26. DOI: https://doi.org/10.1002/9781118358733.wbsyncom040

Sassenhagen, Jona & Alday, Phillip M. 2016. A common misapplication of statistical inference: Nuisance control with null-hypothesis significance tests. *Brain and Language* 162. 42–45. DOI: https://doi.org/10.1016/j.bandl.2016.08.001

Sassenhagen, Jona & Schlesewsky, Matthias & Bornkessel-Schlesewsky, Ina. 2014. The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language* 137. 29–39. DOI: https://doi.org/10.1016/j.bandl.2014.07.010

Schad, Daniel J. & Vasishth, Shravan & Hohenstein, Sven & Kliegl, Reinhold. 2020. How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language* 110. 104038. DOI: https://doi.org/10.1016/j.jml.2019.104038

Scheepers, Christoph & Hemforth, Barbara & Konieczny, Lars. 2000. *Linking syntactic functions with thematic roles: Psych-verbs and the resolution of subject-object ambiguity*. In Hemforth, Barbara & Konieczny, Lars (eds.), *German sentence processing*, 95–135. Dordrecht: Springer Netherlands. DOI: https://doi.org/10.1007/978-94-015-9618-3_4

Schlesewsky, Matthias & Bornkessel, Ina. 2003. Ungrammaticality detection and garden path strength: A commentary on Meng and Bader's (2000) evidence for serial parsing. *Language and Cognitive Processes* 18(3). 299–311. DOI: https://doi.org/10.1080/01690960244000027

Schumacher, Petra B. & Hung, Yu-Chen. 2012. Positional influences on information packaging: Insights from topological fields in German. *Journal of Memory and Language* 67(2). 295–310. DOI: https://doi.org/10.1016/j.jml.2012.05.006

Sorace, Antonella & Keller, Frank. 2005. Gradience in linguistic data. *Lingua* 115(11). 1497–1524. DOI: https://doi.org/10.1016/j.lingua.2004.07.002

Van Valin, Robert D. Jr & LaPolla, Randy J. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139166799

Van de Meerendonk, Nan & Kolk, Herman H.J. & Chwilla, Dorothee J. & Vissers, Constance Th. W. M. 2009. Monitoring in language perception. *Language and Linguistics Compass* 3(5). 1211–1224. DOI: https://doi.org/10.1111/j.1749-818X.2009.00163.x

Van Herten, Marieke & Kolk, Herman H. J. & Chwilla, Dorothee J. 2005. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research* 22(2). 241–255. DOI: https://doi.org/10.1016/j.cogbrainres.2004.09.002

Van Nice, Kathy Y. & Dietrich, Rainer. 2003. Task sensitivity of animacy effects: Evidence from German picture descriptions. *Linguistics* 41(5). 825–850. DOI: https://doi.org/10.1515/ling.2003.027

Wickham, Hadley. 2017. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse.

Wickham, Hadley & Chang, Winston & Henry, Lionel & Pedersen, Thomas L. & Takahashi, Kohske & Wilke, Claus & Woo, Kara & Yutani, Hiroaki & Dunnington, Dewey & RStudio. 2020. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics (Version 3.3.0) [Computer software]. https://CRAN.R-project.org/package=ggplot2.