



How do L2 learners deal with a “dead” language? A psycholinguistic study on sentence processing in Latin

Anna Fiona Weiss 

Received: 25 February 2022 / Revised: 22 February 2023 / Accepted: 11 March 2023
© The Author(s) 2023

Abstract Many decades of research have shown that sentence processing works in an highly incremental and predictive fashion (Marslen-Wilson, 1975) – in the L1 but also (maybe to a lower extent) in the L2. But whereas almost all studies on L2 sentence processing focus on modern languages, it is fully unclear how a language like Latin that fundamentally differs in the way it is taught and used is processed. Thus, the current study focuses on the question if proficient L2 learners of Latin show any evidence for incremental and predictive processing when reading Latin sentences. In a Rapid Serial Visual Presentation (RSVP) task 25 advanced learners of Latin read 384 easy three-word Latin sentences that were manipulated among the factors animacy, argument order and verb position. The results indicate that the participants used the word order and animacy information to incorporate the arguments into the sentence structure on-line and to predict upcoming arguments on the basis of verb information. These findings are interpreted as the first evidence for incremental and predictive processing when reading Latin sentences.

Keywords Latin · Sentence processing · L2 processing · Cue conflicts

Introduction

When native speakers try to extract the meaning of a sentence, they use various cues like word order, case marking or prosody to come to an interpretation of the input. Decades of research have provided convincing evidence that these cues are utilized as soon as they are encountered to establish a coherent representation of the sentence, also known as incremental processing (Frazier, 1987; Marslen-Wilson, 1975). In addition, these cues are used to predict the remaining structure and further arguments of the sentence (Kamide et al., 2003; Kimball, 1975). These routines make comprehension very fast and efficient and allow for a successful communication (Altmann & Kamide, 1999; Federmeier, 2007; Staub & Clifton, 2006).

When people learn a new language, however, they often have difficulties in reaching the same efficient and fast processing routines as native speakers. Many studies show that L2 learners especially fail to utilize grammatical information in real-time comprehension of the L2 (see Kaan, 2014, for an overview). For instance, there is evidence that L2 learners have problems in integrating case-marking information incrementally in languages where the verb is in the final position (Havik et al., 2009; Hopp, 2006; Jackson, 2008). Several explanations for these deficits have been postulated, like effects of the L1, maturational changes because of the age or the reduced input of the L2 (for an overview, see Hopp, 2013). Some

A. F. Weiss (✉)
Catholic University of Eichstätt-Ingolstadt, Eichstätt,
Germany
e-mail: annafiona.weiss@ku.de

authors even argued that L2 speakers construct less detailed syntactic representations than L1 speakers (Shallow Structure Hypothesis; Clahsen & Felser, 2006).

Nonetheless, there is also evidence that L2 speakers are able to process the language incrementally and to use cues (including morpho-syntactic cues) to build a sentence interpretation on-line (see Frenck-Mestre, 2005, and Papadopoulou, 2005, for reviews). But the ability to use morphosyntactic cues and especially to predict upcoming input on the basis of these cues seems to be linked to the extent of overlap between L1 and L2, the type of linguistic structure, and also to the level of proficiency (Mitsugi & MacWhinney, 2016).

However, all of these studies focus on languages that are used for communicative purposes. Learning these languages typically comprises situations in which the learner has to produce or to comprehend the language in real-time. But this is not necessarily true for all languages. Consider Latin, for example. Latin is still learned by thousands of students in school or at university, but it is typically not used for communicative purposes. This also means that language production and real-time comprehension play only a minor role (if at all), since students primarily learn to translate Latin texts into their L1 (Keip & Doepner, 2019; Kuhlmann, 2015, 2019).

This setting, of course, provides a very appropriate testing ground for some important questions with regard to L2 processing. What happens if L2 learners of “dead” languages are exposed to more natural language situations in which it is necessary to comprehend Latin under time pressure and in a clear linear manner (which is typically not the case for Latin)?

One of the most prominent questions in this context is if and how Latin learners make use of linguistic cues during sentence interpretation on-line. Are they able to utilize morphosyntactic information to build sentence structures in real-time and to predict upcoming arguments as observed in other L1 and L2 processing? Or are they fully overstrained and show random or no utilization of cues in real-time because they are not used to this type of natural sentence processing in the context of Latin?

In the current study, this question is addressed by presenting short Latin sentences in a Rapid Serial Visual Presentation (RSVP; Potter, 1984) task to high proficiency learners of Latin with German as L1. This

experimental setting requires the learners to process the Latin sentences very fast and in a linear manner. The sentences are manipulated among the order and animacy of the two arguments and the position of the verb. All sentences are followed by an acceptability judgement, and response times as well as accuracies are measured.

The remaining article is structured as follows. First, findings on cue conflicts and cue additivity in L1 and L2 processing will be summarized and studies on the prediction of thematic roles based on verb information will be presented. After that, the typical settings of learning Latin and the linguistic properties of Latin and German will be outlined as well as current studies on sentence processing in Latin will be introduced. Following that, the rationale of the current study will be defined.

Cue conflicts and cue additivity

The core goal of sentence processing is to figure out who does what to whom. In order to achieve this task, linguistic cues are used. Linguistic cues can be described as mappings between form and function (MacWhinney et al., 1984). The first argument in a sentence (form), for instance, typically points to the actor of the sentence (function; Bader & Bayer, 2006). Typical linguistic cues are (besides word order) case marking, animacy of the argument, plausibility or definiteness (e.g., *the teacher* vs. *a teacher*).

During sentence interpretation, the different linguistic cues are encountered and assessed according to a language-specific cue hierarchy (MacWhinney et al., 1984). This hierarchy of cue strength is based on the reliability and validity of a cue in a certain language.

Consider, for example, the following German sentence:

(1) DenMann liebt die Frau.

The man_{ACC} loves the woman_{NOM/ACC}
The woman loves the man.

In this sentence, the cue word order prefers the first argument “den Mann” as the actor. The cue case marking, however, signals that “den Mann” cannot be the actor because it is marked with accusative case on the determiner. Because case marking is the more reliable cue in German, this leads to the correct interpretation that the woman loves that man. In English, by contrast, not case marking but word order

is the most reliable cue because English does not allow for a flexible word order (due to the absence of case marking). Thus, cue hierarchies are language-specific (MacWhinney et al., 1984).

Studies have shown that if multiple cues converge (i.e., point to the same function), this supports language comprehension. This finding has been reported for adult native speakers, children as well as L2 speakers (Chan et al., 2009; Henry et al., 2017; Henry, Jackson, & Hopp, 2020; Grünloh, Lieven, & Tomasello 2011). On the other hand, if cues conflict with each other (as in the example above), this slows down processing (Boeg Thomsen & Poulsen, 2015; Dittmar et al., 2008; Kim & Sikos, 2011).

An important question in second language acquisition (SLA) research is how learners use these linguistic cues in their L2. Although there is clear evidence that L2 speakers use cues in on-line incremental processing in general, it has been shown that sentence processing in the L2 relies more on lexical-semantic cues (e.g., animacy, plausibility) and surface-level syntactic cues (e.g., word order; Shallow Processing Hypothesis: Clahsen & Felser, 2006). The processing of inflectional morphology (like case marking), by contrast, seems more challenging for the L2 learners (Henry et al., 2020; Hopp, 2015a, 2015b; Mitsugi & MacWhinney, 2016).

Nevertheless, L1 and also L2 speakers make use of linguistic cues during sentence interpretation on-line which for example becomes obvious in the studies on cue additivity and cue conflicts cited above. For the rationale of the current study, this means that if any effects of cue conflicts or cue additivity for Latin learners are observed, it can be concluded that they are able to use linguistic cues on-line. (Just as a reminder: There is little doubt that Latin learners use linguistic cues in general when translating Latin texts, but still the question remains if they are also able to use them under real-time conditions.)

Prediction of selectional restrictions on the basis of verb information

The verb can be described as the center of a sentence because it is the semantic core and provides important information about other entities in the sentence. This information includes at least three aspects. First, the verb is congruent with the subject with regard to number. Second, the verb can require no (intransitive

verbs, e.g., *to sleep*), one (transitive verbs, e.g., *to love*) or two objects (ditransitive verbs, e.g., *to give*). Third, the verb may require an animate subject (e.g., *to love*) or an animate object (e.g., *to hurt*). Thus, the verb contains several cues that allow for predictions about the characteristics of the arguments in the sentence, which is why verb information plays an important role in predictive processing (MacDonald et al., 1994; McRae, Spivey-Knowlton, & Tanenhaus 1998; Trueswell & Tanenhaus, 1994).

In this vein, several studies explored if and how L1 and L2 speakers may utilize verb information to predict upcoming linguistic input. French-Mestre and Pynte (1997), for example, examined how English L2 learners with French as L1 resolved structurally ambiguous sentence structures. They used sentences like.

(2) Every time the dog obeyed the pretty girl showed her approval.

(3) Every time the dog barked the pretty girl showed her approval.

In English, the verb “obey” can either be used transitively or intransitively, which is not the case for the verb “bark” in (3). The eye tracking data showed that both groups were equally garden-pathed when reading English sentences like (2). This indicated that L1 and L2 speakers were able to utilize verb subcategorization information to integrate and predict upcoming information. Similar results were reported by Juffs and Herrington (1995) and Dussias and Cramer (2006). Jackson (2008) also provided evidence that German L2 speakers are able to use the lexical-semantic properties of the thematic verb in wh-questions to rapidly integrate nominal elements into the sentence structures (see also Williams et al., 2001; Altmann & Kamide, 1999; and Hopp, 2015a). In addition, Köhne and Crocker (2010) reported that participants were able to predict further arguments of a sentence on the basis of verb subcategorization information in a language learning paradigm.

As in the L1, the impact of verb information on predictive processing in the L2 obviously depends on the position of the verb. In verb-final structures, it is not possible to use this verb information for predicting upcoming argument structures. Thus, studies show that especially L2 speakers have difficulties in processing sentences where the verb provides the disambiguating information at the end of the sentence.

Table 1 Example stimuli. The stimulus sentences were manipulated among the three factors ARGUMENT ORDER (SO vs. OS), ANIMACY ORDER (AI vs. IA) and VERB POSITION (V1 vs. V2 vs. V3)

Argument order	Animacy	Verb position	
SO	AI	V3	Oratores vinum amant. The orators _{NOM/ACC PL} the wine _{NOM/ACC SG} love _{PL} “The orators love the wine.”
OS	AI	V3	Oratores vinum delectat. The orators _{NOM/ACC PL} the wine _{NOM/ACC SG} pleases _{SG} “The wine pleases the orators.”
SO	IA	V3	Donum sorores delectat. The present _{NOM/ACC SG} the sisters _{NOM/ACC PL} pleases _{SG} “The present pleases the sisters.”
OS	IA	V3	Donum sorores amant. The present _{NOM/ACC SG} the sisters _{NOM/ACC PL} love _{PL} “The sisters love the present.”

SO subject before object, OS object before subject, AI animate before inanimate, IS inanimate before animate, V3 verb last, NOM nominative, ACC accusative, SG singular, PL plural

Havik et al. (2009), for example, investigated how Dutch L2 speakers with German as L1 processed subject-object ambiguities in relative clauses. Although both structures are highly comparable in Dutch and German (in both languages, the verb is in the final position), Havik and colleagues did not observe a subject-before-object preference for all L2 speakers in on-line processing. This means, the L2 speakers did not show the same processing routines as Dutch native speakers did.

Learning Latin

Although there are no native and only a very few fluent speakers of Latin anymore, this language is still learned by thousands of students in school or at university, especially in Germany, Austria, Italy and the UK. The way Latin is taught and learned, however, differs clearly from the acquisition of other languages. In the following, some important differences that highlight the special status of Latin will be summarized. Since Latin learners with German as L1 are examined, the focus of this summary is on the situation in Germany.

The most important characteristics of learning Latin are:

- (a) Overall, the learning process is highly controlled since there is very little exposure to Latin outside of school or university contexts.

- (b) Latin is usually taught in the learners' L1 with a strong focus on formal grammar instruction.
- (c) There is a strong preference for the visual modality. Students are very rarely exposed to oral Latin language, except for reading aloud the Latin texts.
- (d) Translation from Latin into the L1 forms an integral part of Latin language learning (Keip & Doepner, 2019; Kuhlmann, 2015, 2019) and it often constitutes the predominant way of dealing with Latin input.
- (e) There is a strong preference for language comprehension, whereas language production only plays a minor role. Only when studying Latin at university, production tasks become more important since all students have to take courses on translating sentences from their L1 into Latin.
- (f) Even highly proficient L2 learners often have difficulty in reading Latin texts for comprehension. It has to be acknowledged that this has not been empirically explored as yet, but it appears that most of the university students still apply translation strategies and need more time and further assistance (for instance in the form of vocabulary) to derive the meaning of original Latin texts compared to proficient L2 speakers of other languages.

Table 2 Accuracy results (in percent) and reaction times (in milliseconds) for each condition separately

Example stimuli	Literal translation	Argument order	Animacy order	Verb-position	Accuracy (%)	RT (ms)
Amant oratores vinum	<i>The orators love the wine</i>	SO	AI	1	94.78 (8.66)	1281 (616)
Delectat oratores vinum	<i>The wine pleases the orators</i>	OS	AI	1	85.95 (16.47)	1798 (790)
Delectat donum sorores	<i>The present pleases the sisters</i>	SO	IA	1	86.80 (17.59)	1911 (673)
Amant donum sorores	<i>The sisters love the present</i>	OS	IA	1	94.95 (10.14)	1442 (866)
Oratores amant vinum	<i>The orators love the wine</i>	SO	AI	2	97.28 (5.27)	1080 (627)
Oratores delectat vinum	<i>The wine pleases the orators</i>	OS	AI	2	71.97 (20.27)	2054 (696)
Donum delectat sorores	<i>The present pleases the sisters</i>	SO	IA	2	88.90 (13.43)	1599 (787)
Donom amant sorores	<i>The sisters love the present</i>	OS	IA	2	88.54 (12.86)	1490 (594)
Oratores vinum amant	<i>The orators love the wine</i>	SO	AI	3	94.41 (7.54)	1431 (641)
Oratores vinum delectat	<i>The wine pleases the orators</i>	OS	AI	3	76.16 (23.13)	2206 (625)
Donum sorores delectat	<i>The present pleases the sisters</i>	SO	IA	3	84.99 (19.07)	1946 (789)
Donum sorores amant	<i>The sisters love the present</i>	OS	IA	3	90.76 (10.13)	1958 (631)

Standard deviations are given in parentheses

RT reaction time, *ms* milliseconds, *SO* subject before object, *OS* object before subject, *AI* animate before inanimate, *IA* inanimate before animate

(g) Since Latin is not used for communicative purposes nowadays (some exceptions provided), there is no need to process Latin in a fast manner. Therefore, Latin learners are typically never exposed to situations in which they have to process Latin in real-time.

These characteristics distinguish Latin from other L2 learning contexts and stress its significance for L2 sentence processing studies. Thus, Latin allows us to explore questions that can hardly be examined by using other languages. Especially the fact that learners typically translate Latin into their L1 without any time pressure, may have important implications for sentence processing because it may shed light on the question whether and how incremental and predictive processing routines can be transferred from one language to another. More specifically, it may be

asked if Latin learners are able to utilize morpho-syntactic or lexical-semantic cues on-line in the L2 when they are exposed to situations in which incremental and predictive processing of Latin is required, even though they have little or no experience with processing Latin in real-time.

If it was possible to find evidence that these strategies can be transferred to Latin that would also have implications for learning and teaching Latin: It would suggest that it may be a fruitful approach to treat Latin more like a “natural” language and to provide students with exercises that explicitly practice incremental processing routines.

Please note that in contrast to modern languages it is not possible, of course, to compare processing of Latin L2 learners with processing of native speakers. When conducting studies with Latin L2 learners in the context of sentence processing, the rationale is just to

Table 3 Results of the generalized linear mixed-effects models analysis for accuracy

model_acc_max = glmer(accuracy ~ animacy_order*argument_order*verb_position + (1 + animacy_order*argument_order*verb_position|subj) + (1 + animacy_order*argument_order*verb_position|item), data = data, family = "binomial", control = glmerControl(optimizer = "nlopt", optCtrl = list(maxfun = 1e5), calc.derivs = FALSE))

Fixed effects	Estimate	SE	z value
Intercept	2.90	0.29	10.09
Animacy_order	0.19	0.22	0.86
Argument_order	- 0.55	0.17	3.21
Verb_position 2	0.05	0.16	0.29
Verb_position 3	- 0.47	0.16	2.88
Animacy_order x Argument_order	- 0.84	0.17	4.79
Animacy_order x Verb_position 2	0.24	0.24	1.03
Animacy_order x Verb_position 3	- 0.11	0.21	0.55
Argument_order x Verb_position 2	- 0.64	0.23	2.72
Argument_order x Verb_position 3	0.12	0.19	0.62
Animacy_order x Argument_order x Verb_position 2	- 0.30	0.24	1.22
Animacy_order x Argument_order x Verb_position 3	0.26	0.17	1.56

The full model is given in the first row. Reference (omitted) level is "verb position 1"

look for processing routines that have been shown to be typical for L1 and L2 processing in general.

Linguistic properties of latin and German

Latin has a rich inflectional system. All nouns are case-marked but not necessarily unambiguously. For example, for neuter nouns like *templum* (the temple) the ending *-um* marks nominative as well as accusative case.

The rich inflectional system allows for flexible word order, and especially in poetry texts, the word order can be fully scrambled. This is one reason why learning Latin is commonly considered as particularly challenging. In order to face these challenges, teachers often provide their students with certain translation methods (for an overview, see Keip & Doepner, 2019). These methods partly emphasize processing strategies that are very uncommon in natural reading, for example the "construction method". Here students are instructed to firstly search for the verb and then to identify its dependent constituents.

German also allows for a flexible word order, but with greater restrictions than in Latin. As in Latin, all three verb positions are possible but only in certain sentence structures. In yes–no questions, on the one hand, the verb appears in the first position like.

(4) Liebt der Mann die Frau?

Does the man love the woman?

In all main clauses, on the other hand, the verb has to be in the second position:

(5) Der Mann liebt die Frau.

The man loves the woman.

And finally, all subclauses require a verb in the last position:

(6) Alle wissen, dass der Mann die Frau liebt.

All know that the man loves the woman.

Within these restrictions, the order of subject and object(s) can be scrambled. However, there are many studies showing that native speakers of German tend to interpret the first argument as the subject and revise this interpretation as soon as conflicting evidence is encountered (Hemforth et al., 1993; Schlesewsky et al., 2000).

In German, all nouns are case-marked but typically on the determiner. As in Latin, these markings can be ambiguous, for example for feminine nouns like *die Frau*. Here the determiner *die* is ambiguous with regard to nominative and accusative case. Thus, case marking and word order are not fully reliable cues neither in German nor in Latin.

Taken together, Latin and German share many linguistic properties, which is why German L1 speakers provide an appropriate participant group for the current experiment. If it is the objective to find any evidence for cue utilization in on-line Latin

processing, it is most likely to find them for L1 speakers of linguistically similar languages (Ellis, 2006; MacWhinney, 1992; see also Jackson & Dussias, 2009).

Studies on sentence processing in Latin and the present study

Latin has been subject to psycholinguistic research only in a few cases. In most of these studies, Latin was used as a tool to investigate general questions of language teaching or language processing. Ellis and Sagarra, for example, used Latin to study associative learning by teaching temporal reference to university students who never learned Latin before (Ellis & Sagarra, 2010, 2011). Within *The Latin Project* (Stafford et al., 2012), on the other hand, Latin was used as a new L2 in order to explore how different instructional treatments (e.g., explicit grammar practice, explicit feedback) affect the learning process. This design was also used to examine the impact of different L2 languages on learning Latin as a L3 (Sanz et al., 2015). More recently, VanPatten and Smith (2019) examined how word order affects the acquisition of case marking by using simple SOV (subject–object–verb) and SVO (subject–verb–object) sentences in Latin.

All of these studies focus on very short sentences in Latin, typically consisting of 3–4 words. This is done – inter alia – because it is hard to find participants that are able to process more complex Latin sentences in a more or less automatized manner as necessary for psycholinguistic studies. More importantly, these short sentences allow for the very controlled manipulation of different experimental (linguistic) factors. Stafford and colleagues (2012), for example, examined the impact of the factors word order, verb agreement and case marking on the processing of Latin sentences. For this, they adopted an experimental design often used in the context of the Competition Model (MacWhinney et al., 1989, 2002). In the current study, this idea is adopted and also short Latin sentences consisting of two nouns and one verb are used.

The current study focuses on three cues and their impact on sentence processing: order of the two arguments, animacy of the two arguments and position of the verb. All of these cues have been subject to various studies of sentence processing in the L1 as

well as in the L2, in particular in the context of the Competition Model (for an overview, see MacWhinney, 2005).

Given that this study is the first one focusing on sentence processing in Latin instead of learning Latin as a new language, the rationale is to look for any evidence of cue usage during on-line processing. Since proficiency has been shown to have a strong impact on the morpho-syntactic processing in the L2 (Hopp, 2015a; Jackson, 2008), the study focuses on Latin learners that are highly proficient. The mostly comparable group (e.g., with regard to age, level of education and language proficiency) of highly proficient Latin learners are university students, studying Latin as a major subject. This is why these students were chosen as participants.

In addition, the decision was made to use a Rapid Serial Visual Presentation (RSVP) paradigm for presenting the sentences (Potter, 1984). In this paradigm, the sentence is presented automatically word by word on a computer screen in a very rapid manner (typically 250–600 ms each). After the presentation, the subject is asked to answer a question about the sentence as fast and accurately as possible. The reaction time and accuracy results are measured. The rationale behind this method is that the structure of the sentence presented (i.e., the manipulation) should affect the accuracy and reaction times. Since the response of the subject is only measured after presenting the whole sentence, this method is not informative with regard to where exactly in the sentence processing difficulties appeared – in contrast to clear on-line methods like eye tracking.

The rationale why a RSVP paradigm is still used is that this study targets on-line sentence processing, which is why the participants have to be put under time pressure. When using self-paced reading instead, which would provide a more fine-grained picture of the processes conducted online, the participants would not have the time pressure as required by the research question. For reading studies using eye tracking, on the other hand, the Latin sentences are not long enough and have too little variation. Visual world paradigms, finally, would require auditive sentence presentation, which clearly increases the difficulty of the task given that Latin learners typically are not exposed to oral Latin language. Thus, a RSVP task is used knowing that the results just give a first impression and have to

be replicated and examined in more detail by following research.

In the current design, the participants are asked for an acceptability judgement after each sentence while the reaction time and accuracy of this judgement are measured. Acceptability includes grammatical as well as semantic aspects, which is why this seems an appropriate possibility to measure the comprehension of the sentences. As mentioned above, the RSVP task only provides after-sentence measures which are only partly informative with regard to processing mechanisms applied on-line.

The hypotheses are as follows:

- (a) It is expected that if multiple cues converge, sentence processing is faster and more accurate. If cues conflict, this slows down reaction time and reduces accuracy.
- (b) Verb-initial sentences are expected to be processed faster and more accurately since verb information allows for predictions on the number and animacy status of the subject.

Method

Participants

Twenty-five students from the University of Marburg were tested (14 female; mean age: 24 years, ranging from 20 to 29 years). All participants were native speakers of German and highly proficient learners of Latin. L2 acquisition started at secondary school and participants were enrolled in major studies in Latin (as future teachers). They had already finished undergraduate studies. At this level, students are typically able to translate (not to read) unknown original Latin texts (like Caesar, Cicero, Ovid) but with some time of preparation and the help of a dictionary. These participants were chosen as it was the most proficient group that was available for examination. Participants were naïve concerning the purpose of the study and received 7 Euro for participation.

Materials and design

Each stimulus sentence consisted of three Latin words (one animate plural noun, one inanimate singular noun and one verb, in varying word order; see Table 1).

Both nouns were ambiguously case-marked as subject and object, so that only the verb contained the disambiguating number information. Also, the verb either required an animate subject (as *amare* [to love]) or also allowed for an inanimate subject (as *delectare* [to please]).

In our experiment the factors ARGUMENT ORDER (subject before object [SO] vs. object before subject [OS]), ANIMACY ORDER (animate before inanimate [AI] vs. inanimate before animate [IA]) and VERB POSITION (verb first [V1] vs. verb second [V2] vs. verb last [V3]) were manipulated, leading to a 2x2x3 experimental design (see Table 1).

Four sentences for each of the 12 conditions were created. The words were taken from a set of only 24 concrete and high-frequent Latin words from the core vocabulary, which were allocated to four sets. In order to maximize the experimental power, the resulting 48 experimental sentences were doubled (i.e., the same sentences were repeated), leading to 96 stimulus sentences in total. All stimuli were grammatical and plausible sentences of Latin (see Appendix for a table of all stimuli).

The procedure to double the stimuli sentence and not to run different lists according to a Latin square design, for instance, is unusual. Nevertheless, this procedure was chosen for the following reasons:

- a) The current study primarily focuses on semantic and syntactic processing and thus it was necessary to ensure that the participants had no difficulties with vocabulary. Although the participants were all advanced learners of Latin, they had very little experience with recalling a word meaning under time pressure. Thus, the number of Latin words had to kept low.
- b) The stimuli sentences were very short and the semantic content was low. In addition, they were interspersed with a high number of filler sentences (see below). It was therefore concluded that the repetition effect was negligible.
- c) All of the participants in the pilot study and in the experiment reported that the task was very demanding and pretty unusual for them in the context of Latin. No one stated that he or she had realized that stimulus sentences appeared twice in the experiment.

In addition to the stimulus sentences, 288 filler sentences were created in total (using the same lexical

material described above) to prevent participants from certain reading strategies. 96 filler sentences were created by replacing the verb-type in the plausible stimulus sentences (e.g., *amant* vs. *delectant*), leading to implausibility. The remaining 196 fillers included ungrammatical sentences, sentences with unambiguous case marking on the animate argument (nominative or accusative singular) as well as ambiguous plural forms of the inanimate argument (nominative and accusative). In sum, there was an even number of plausible and implausible sentences in the experiment and each participant read all 384 sentences but in random order.

Procedure

All 384 sentences were presented word-by-word on a computer screen (for 500 ms each word) as a Rapid Serial Visual Presentation (RSVP) task using the software PRESENTATION (version 19.0, Neurobehavioral Systems, 2016). Before the sentence started, a fixation star in the middle of the screen appeared for 400 ms. In addition, after each trial the participants had to indicate whether the sentence was acceptable or not by pressing the “F” (= NO) and “J” (= YES) buttons on the keyboard. The participants were asked to answer as fast and accurately as possible, and the type of response as well as the reaction time were recorded. Probing acceptability has the advantage that it includes grammaticality and plausibility judgments. In addition, the linguistic information of the question itself is very low, especially in contrast to questions that contain evaluations of a statement about the sentence. Thus, it seems to be an appropriate approach to examine the comprehension of the experimental sentences (Myers, 2017).

The experiment consisted of 8 blocks (with 48 sentences each) and took about 30 min. Between the blocks the participants were allowed to have a short break. In order to make the subjects familiar with the procedure, the experiments started with a practice session containing 20 sentences.

Given the high degree of polysemy in Latin, it was necessary to ensure that the required word meaning could be immediately accessed during sentence processing (see above). Hence, in preparation for the experiment, participants were given 24 vocabulary items with one German translation each. These were tested prior to the start of the experiment. Only

subjects who passed this vocabulary test without any mistakes were allowed to participate in the study (all subjects passed).

In addition, the participants filled in a short questionnaire about their language biography and gave written informed consent for participating in the study.

Results

Prior to analysis, participants with a mean accuracy of less than 70% were excluded (this affected the data of two participants). In addition, trials with latencies of 2.5 SD above group mean were excluded (< 2% of all trials). The remaining 23 participants reached a mean accuracy level of 87.81% (SD: 32.71; range 70.00–98.70%) and took 1901 ms on average to respond (SD: 1514). The accuracy results and reaction times for each condition separately are provided in Table 2. Note that only correct trials entered reaction time analysis.

In order to check for significant differences between the stimulus conditions, generalized linear mixed-effects models for accuracy and linear mixed-effects models for reaction times were fitted, using the lme4 package (Bates et al., 2014) and the NLOpt nonlinear-optimization package for R (<http://github.com/stevengj/nlopt>). Following Barr et al. (2013), the maximal random effect structure justified by the experimental design was used. Both two-level predictors (animacy order and argument order) were encoded as sum contrasts with zero as the mean of the two levels (-1,1). The three-level predictor verb position was also encoded as a sum contrast with “verb position 1” being the reference level. In this case, the intercept reflects the grand mean and each effect reflects whether the level is reliably different from the grand mean (Brehm & Alday, 2022). Because reaction times violate the assumption of normally distributed variables, they were converted to inverse reaction times, according to Brysbaert and Stevens (2018). Following convention, $|z|$ and $|t| > 2$ were treated as significant.

Accuracy results

The exact formula of the model as well as the full results are given in Table 3.

Fig. 1 Main effects of ARGUMENT_ORDER and VERB_POSITION 3 as well as the interaction of VERB_POSITION 2 and ARGUMENT_ORDER for accuracy results.

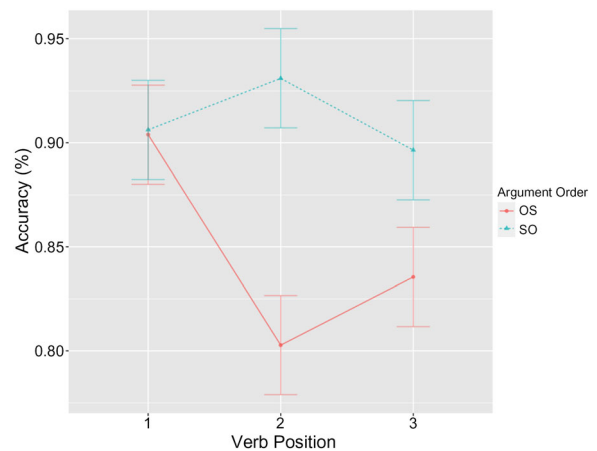
Abbreviations: *OS* object before subject, *SO* subject before object, *AI* animate before inanimate, *IA* inanimate before animate. The bars indicate standard deviation

In the analysis the main effect of ARGUMENT_ORDER ($|z|= 3.21$) was significant, showing that sentences with an object-before-subject structure were generally processed less accurately than sentences with the reversed argument order (OS: 84.74% vs. SO: 91.14%).

There was also a significant main effect of VERB_POSITION 3 ($|z|= 2.88$, see Fig. 1). This effect accounted for the fact that stimuli where the verb was in the final position were processed less accurately (86.54%) compared to the mean of all three verb types (87.97%). Further inspection with linear hypothesis testing revealed that the accuracy for sentences with V1 also significantly differed from mean ($X^2(1, n = 2) = 5.88, p < 0.05$) but in a positive direction (90.57%). Although the mean of accuracy for V2 sentences was numerically comparable to V3 sentences (86.54%), there was no main effect for V2

Fig. 2 Interaction of ARGUMENT_ORDER and ANIMACY_ORDER for accuracy results.

Abbreviations: *OS* object before subject, *SO* subject before object, *AI* animate before inanimate, *IA* inanimate before animate. The bars indicate standard deviation



sentences ($|z|= 0.29$), probably due to the significant interaction with ARGUMENT_ORDER (see below).

Also, the interaction of ANIMACY_ORDER x ARGUMENT_ORDER was significant ($|z|= 4.79$, see Fig. 2). In order to check this interaction in more detail, again general linear mixed-effects models were fitted for SO and OS sentences separately (in the way described above but using only animacy order as a predictor). For both sentence types there was a main effect of ANIMACY_ORDER (SO: Estimate 0.64, SE 0.20, $|z|= 3.17$; OS: Estimate -0.51, SE 0.14, $|z|= 3.63$): sentences with animate subjects were always processed more accurately than those with inanimate subjects across word order. This also means that for SO word order, AI order was more accurately processed than IA (95.50% vs. 85.84%) but for OS word order, IA was processed more accurately than AI (91.41% vs. 78.06%).

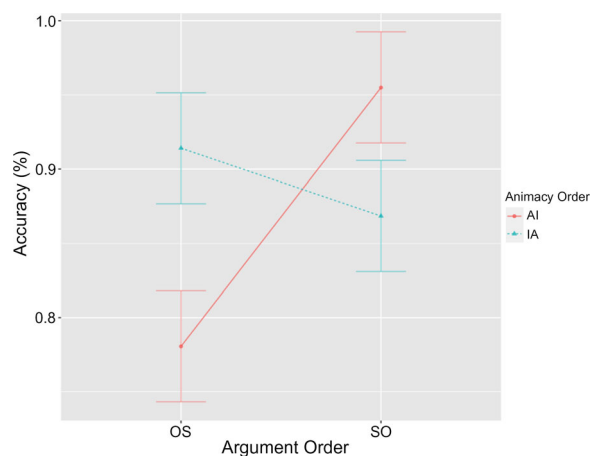


Table 4 Results of the linear mixed-effects models analysis for reaction time

Model_RT_max = lmer(inv_reactiontime ~ animacy_order*argument_order*verb_position + (1 + animacy_order*argument_order*verb_position|subj) + (1 + animacy_order*argument_order*verb_position |item), data = data, control = lmerControl(optimizer = "nlopt", optCtrl = list(maxfun = 1e5), calc.derivs = FALSE))

Fixed effects	Estimate	SE	t value
Intercept	- 0.08	0.10	0.83
Animacy_order	- 0.01	0.05	0.20
Argument_order	0.11	0.05	2.04
Verb_position 2	- 0.10	0.06	1.69
Verb_position 3	0.21	0.07	2.99
Animacy_order x Argument_order	0.15	0.04	3.58
Animacy_order x Verb_position 2	- 0.04	0.05	0.81
Animacy_order x Verb_position 3	- 0.01	0.06	0.22
Argument_order x Verb_position 2	0.13	0.07	1.82
Argument_order x Verb_position 3	- 0.01	0.06	0.19
Animacy_order x Argument_order x Verb_position 2	0.01	0.06	0.12
Animacy_order x Argument_order x Verb_position 3	- 0.08	0.05	1.54

The full model is given in the first row. Reference (omitted) level was “verb position 1”

Lastly, the interaction of ARGUMENT_ORDER x VERB_POSITION 2 was significant ($|z|= 2.72$, see Fig. 1). When the verb was in the second position and the object preceded the subject, the accuracy dropped from 93.10% (SO sentences) to 80.27% (OS sentences). In order to also check the omitted group (VERB_POSITION 1) for this interaction, a general linear mixed-effects model was fitted for V1 sentences separately (in the way described above but using only word order as a predictor). As expected, there was no main effect of argument order for V1 sentences (Estimate -0.19, SE 0.16, $|z|= 1.21$).

Reaction time results

The exact formula of the model as well as the full results are given in Table 4.

In the analysis the main effect of ARGUMENT_ORDER ($|t|= 2.04$) also was significant, showing that subject-first sentences were processed faster than object-first sentences (OS: 1796 ms vs. SO: 1518 ms).

In addition, there was again a significant main effect of VERB_POSITION 3 ($|t|= 2.99$, see Fig. 3). When the verb was in the final position, reaction times significantly increased compared to group mean (V3: 1845 ms, group mean: 1636 ms). In contrast to the accuracy results, there was only a small difference

between V1 and V2 sentences (1598 ms vs. 1512 ms) and linear hypothesis testing revealed that the main effect for V1 sentences was only marginally significant ($X^2(1, n = 2) = 3.68, p = 0.05$). In addition, ARGUMENT_ORDER did not significantly interact with VERB_POSITION 2 ($|t|= 1.82$) as was the case for the accuracy results.

But as for the accuracy results, a significant interaction of ANIMACY_ORDER x ARGUMENT_ORDER was observed ($|t|= 3.58$, see Fig. 3) and as for accuracy this interaction was resolved towards animacy order by fitting separate linear mixed-effects models. The results revealed again significant main effects of ANIMACY_ORDER for both SO and OS sentences (SO: Estimate -0.16, SE 0.06, $|t|= 2.56$; OS: Estimate 0.13, SE 0.05, $|t|= 2.69$), representing the same picture as for accuracy: the reaction times for sentences with animate subjects were always shorter than for those with inanimate subjects across word order. This means accordingly that for SO word order, the reaction times for sentences with AI order were shorter than for sentences with IA order (1269 ms vs. 1815 ms) but for OS word order, participants responded faster to sentences with IA than with AI order (1636 ms vs. 1993 ms) (Fig. 4).

Fig. 3 Main effects for ARGUMENT_ORDER and VERB_POSITION 3 for reaction times. Interactions and main effect for VERB_POSITION_2 did not reach significance. Abbreviations: *OS* object before subject, *SO* subject before object, *AI* animate before inanimate, *IA* inanimate before inanimate. The bars indicate standard deviation

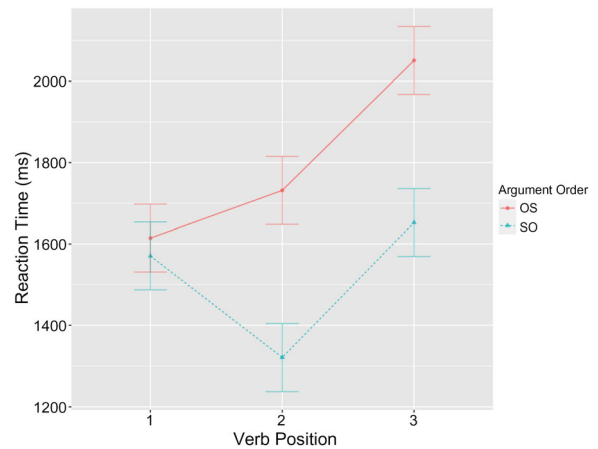
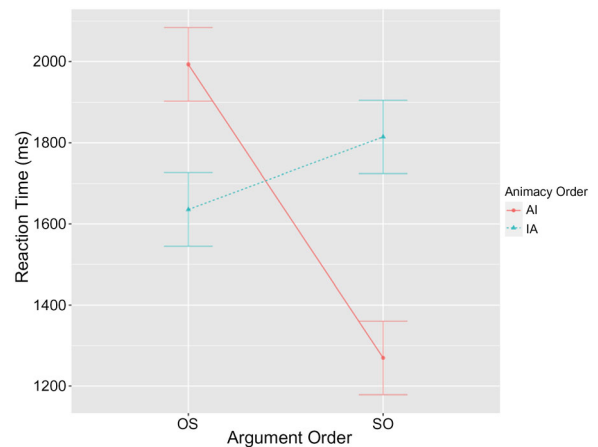


Fig. 4 Interaction of ARGUMENT_ORDER and ANIMACY_ORDER for reaction times. Abbreviations: *OS* object before subject, *SO* subject before object, *AI* animate before inanimate, *IA* inanimate before inanimate. The bars indicate standard deviation



Discussion

The rationale of the current study was to look for evidence for incremental and predictive processing of Latin sentences by L2 learners. In order to investigate this, proficient learners of Latin were asked to read easy three-word sentences in Latin that were manipulated among the factors argument order, animacy order and verb position, and to decide as fast as possible whether these sentences were acceptable. The first hypothesis was that if multiple cues converge, sentence processing is expected to be faster and more accurate. If cues conflict, on the other hand, this slows down reaction time and reduces accuracy. The second hypothesis states that verb-initial sentences are expected to be processed faster and more accurately since verb information allows for predictions on the number and animacy status of the subject.

Overall, the span of the accuracy results (70.00–98.70%) indicates that the task was challenging but practicable for the participants. This is an important point given that the experimental design, in particular the time pressure, was very unfamiliar to the participants. Obviously, they were nevertheless able to process easy Latin sentences under real-time conditions and came (for the most part) to correct interpretations of the sentences. Accordingly, the results seem to be reliable indications on how the Latin learners did process these sentences.

The mixed-models analysis revealed a significant main effect for argument order and a clear interaction of animacy order and argument order in both reaction times and accuracy. In addition, a significant main effect for V3 sentences in both measures was observed. For accuracy, the argument order further significantly affected the processing of V2 sentences.

In the following, these results are discussed in more detail.

Processing animacy order and argument order

First of all, the main effect of argument order as well as the interaction of animacy order and argument order provide evidence that the participants did use these cues for sentence processing.

The main effect of argument order indicates that subject-first sentences were processed faster and more accurately than object-first sentences. Thus, it seems that participants used the cue argument order and interpreted the first argument (which was ambiguously case-marked) as the subject. In cases this interpretation turned out to be wrong (especially when the verb was encountered) a reanalysis had to be carried out which caused increased reaction times. The accuracy results also show that in about 15% of cases the participants did not reach the correct interpretation at all.

In addition, the results can be interpreted as an indication for cue additivity and cue conflicts: In cases in which both cues (animacy order and argument order) converged, processing was fastest and most accurate. This was true for SO structures where the animate argument preceded the inanimate argument as well as for OS structures in which the inanimate argument preceded the animate argument. In both sentence types, both cues favored the same argument as the actor. In cases, however, in which the two cues favored different arguments as the actor (SO-IA / OS-AI sentences), processing speed was reduced and accuracy dropped down. Thus, these results mirror well established findings of L1 and L2 sentence processing, especially in German (Brandt et al., 2016; Chan et al., 2009; Dröge et al., 2020; Jackson & Roberts, 2010; MacWhinney et al., 1984).

Given that there was no control group of native Latin speakers, it is not clear whether these strategies are just transfer strategies from the L1 or particular strategies of the L2. There is broad evidence that cue strength in the L1 is transferred to the L2, especially for novice learners (for overviews, see MacWhinney, 2002, 2005). Thus, in the beginning learners seem to transfer cue hierarchies from their L1 into their L2, but when learning proceeds, they adjust more and more to the cue hierarchies in the L2 (based on linguistic evidence and instruction) to become more native-like.

Since the two cues under consideration are similarly reliable in both Latin and German, the Latin learners may just have applied their L1 strategies. Especially the fact that there is broad evidence for a strong subject-before-object preference in German (Hemforth et al., 1993; Schlesewsky et al., 2000) which was also found in the Latin data supports this hypothesis. Interestingly, this preference was even present in the interaction with animacy order, indicating that the cue animacy was not able to fully overwrite this preference.

However, the question of cue transfer cannot be answered on the basis of the current data yet. It therefore seems to be fruitful to examine cue interpretation in speakers of other L1s like English, which show distinct cue hierarchies from Latin (MacWhinney et al., 1984).

In addition, argument order and animacy were examined, but the cue case marking was not. Thus, there was evidence for using lexical-semantic (animacy) and syntactic cues (argument order) during on-line sentence processing in Latin, but the core morpho-syntactic cues were not examined. Since several studies have shown that L2 learners in particular fail to utilize morpho-syntactic information for predictions (Clahsen & Felser, 2006), examining case processing in Latin should also be a subject for future research.

Processing verb information

The data revealed that processing speed was reduced and accuracy dropped down when the verb was in the last position. In addition, there was a pronounced effect of argument order in V2 sentences for accuracy. In the following, these results will be discussed, starting with V1 sentences.

When the verb was in the first position, there was no difference between SO and OS structures. Since the verb allowed for the prediction of further arguments due to number agreement (one noun was singular and the other plural) and due to the thematic roles (some verbs required animate subjects, others did not), this clearly indicates that participants were able to use this verb information for sentence processing on-line and to predict the characteristics of the further arguments.

These results are in line with other research on processing verb information (e.g., Köhne & Crocker,

2010) but the strength of the prediction effect is still somewhat surprising. For example, Hopp (2020) concludes from his experiment with German L1 and L2 speakers, that “adult L2 learners routinely do not make morphosyntactic predictions in the first place” (p. 642). Maybe the clear prediction effects in the current experiment point to the special status of the verb in learning Latin. In popular Latin translation methods like the “construction method”, for instance, the students are directly encouraged to derive predictions from the verb. Since Hopp (2020) further reports in experiment 2 that L2 learners can learn to derive predictions after explicit exposure, this may be a plausible explanation.

When the verb was in the second position, however, argument order significantly affected sentence processing. SVO sentences were clearly processed faster and more accurately than OVS sentences. Although more fine-grained measures are needed here, these results seem to indicate that participants typically interpreted the first argument as the subject. In cases where new information during sentence processing questioned this interpretation (namely, when the verb was encountered which was not compatible with this supposed subject because of animacy or number disagreement), this required reanalysis. Accuracy data revealed that for about 20% of these OVS sentences, participants did not come up with the correct interpretation since the wrong interpretation was either not detected or participants failed to revise their initially adopted structure.

When the verb was in the final position, accuracy dropped down and participants needed longer to respond. This fits well to the consideration that the verb contains the relevant information for disambiguation as discussed above. Thus, reduced accuracy may indicate that participants had trouble in re-analyzing the sentence when finally encountering the verb and / or that they had difficulty in remembering the two arguments encountered before. The longer reaction times may also reflect reanalysis but in addition they may reflect evaluating processes at the end of the sentence before making a choice and pressing the answer button (see Weiss, 2020, for similar observations). However, there was no significant interaction with argument order. Thus, it cannot clearly be determined how the order of the two arguments affected the reaction times and accuracy results of V3 sentences.

In sum, these effects seem to confirm the second hypothesis: Verb information was used by the participants to predict further arguments of the sentence. On the other hand, it seems that the participants had nonetheless difficulties in using the verb information for revising initially wrong interpretations, indicated by the reduced accuracy results for V2 and V3 sentences.

On-line processing of cues in Latin and implications for further research

The current experiment was focusing on the question of how highly proficient Latin learners use different cues for sentence interpretation in real-time although the participants are not used to processing Latin in this fashion. The results indicate that Latin learners were indeed able to utilize argument order and animacy information to identify the actor of the sentence. In addition, there was evidence that participants were able to predict further arguments on the basis of verb information.

These results suggest, for the first time, that Latin can be processed similarly to other L2s despite its very different way of acquisition. This finding has important implications for our understanding of L2 learning because it shows that general processing strategies of the L1 can be transferred to the L2 without further practice. In addition, the findings may also have implications for learning Latin because they suggest that Latin in principle can be processed like any other modern language. This insight might fundamentally change the way Latin is taught nowadays. Taking this finding seriously, it could be suggested that Latin teaching should focus more on natural reading than on formal grammar instruction and translation, treating Latin more like a modern language.

However, the results of our experiment also have to be interpreted with caution. First, it has to be acknowledged that using a RSVP task provides only limited insights into the processes conducted on-line. Thus, more fine-grained measures are needed to exactly determine which and how cue information is used during on-line processing of Latin by L2 learners.

Second, only very easy Latin sentences were used that are not comparable to the sentences normally encountered when learning Latin. Moreover, the participants were proficient learners of Latin. It is

unclear, therefore, if incremental processing can be also observed when students at school read Latin sentences. In any case, it seems to be a very promising way to train students in incremental processing (maybe even by using training sessions like the experimental design described above) and see how this affects Latin sentence processing.

Another important issue for future research seems to be the impact of the L1. The results suggest that the L1 shaped the way Latin sentences are processed but it was not possible to provide clear evidence for this claim. Given that Latin and German are similar in several respects (e.g., flexible word order, case marking) it would be particularly interesting to examine the influence of the L1 more systematically. A fruitful approach could be to do the same experiment with students from other countries, for example English speaking countries, because English has a very strict word order. This would allow us to understand in more detail which processing strategies are transferred from the L1 and which are genuinely Latin. Note that we do not have native speakers of Latin anymore whose data would be needed to identify the “original” strategies.

Summary and conclusion

The current experiment provides the first quantitative study on processing Latin using a very controlled experimental design. The main question was whether it proves possible to find any evidence that L2 learners of Latin process Latin sentences incrementally and derive predictions on the basis of verb information. The results show for the first time that proficient Latin learners are indeed able to use language information cues like argument order and animacy to identify the actor of a sentence and to predict upcoming arguments during on-line processing. These findings are interpreted as clear evidence for incremental and predictive processing.

However, the RSVP paradigm just provides a first impression, more fine-grained measures are needed to derive reliable conclusion. It is further still unclear if the applied processing strategies are part of the learners’ L1 (which was German) or if the learners used Latin-specific strategies. Also, more research is needed to see whether these results can be generalized

about other populations (e.g., Latin learners at school) or other language material (e.g., more complex texts).

Acknowledgments The author thanks Delaram Bextermöller for help with experimental set-up and data collection. The author also would like to thank an anonymous Reviewer for his or her very helpful comments on earlier versions on the manuscript as well as Thomas H. Kaal for proofreading the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The project was partly funded by “Hessischer Altphilologenverband” (Classical scholar association of Hestia).

Declarations

Conflict of interest The author has no conflicts of interest to declare that are relevant to the content of this article. The funding institution had no influence on experimental design, data collection, data processing or data interpretation.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

The following table contains all stimulus sentences used in the experiment. Literal translations are given in the third column.

Set	Example stimuli	Literal translation	Argument order	Animacy order	Verb position
1	Amant oratores vinum	<i>The orators love the wine</i>	SO	AI	1
1	Delectat oratores vinum	<i>The wine pleases the orators</i>	OS	AI	1
1	Delectat donum sorores	<i>The present pleases the sisters</i>	SO	IA	1
1	Amant donum sorores	<i>The sisters love the present</i>	OS	IA	1
1	Oratores amant vinum	<i>The orators love the wine</i>	SO	AI	2
1	Oratores delectat vinum	<i>The wine pleases the orators</i>	OS	AI	2
1	Donum delectat sorores	<i>The present pleases the sisters</i>	SO	IA	2
1	Donom amant sorores	<i>The sisters love the present</i>	OS	IA	2
1	Oratores vinum amant	<i>The orators love the wine</i>	SO	AI	3
1	Oratores vinum delectat	<i>The wine pleases the orators</i>	OS	AI	3
1	Donum sorores delectat	<i>The present pleases the sisters</i>	SO	IA	3
1	Donum sorores amant	<i>The sisters love the present</i>	OS	IA	3
2	Portant victores ferrum	<i>The winners bear a sword</i>	SO	AI	1
2	Vulnerat victores ferrum	<i>The sword injures the winners</i>	OS	AI	1
2	Portant telum fratres	<i>The brothers bear a weapon</i>	SO	IA	1
2	Vulnerat telum fratres	<i>The weapon injures the brothers</i>	OS	IA	1
2	Victores portant ferrum	<i>The winners bear a sword</i>	SO	AI	2
2	Victores vulnerat ferrum	<i>The sword injures the winners</i>	OS	AI	2
2	Telum portant fratres	<i>The brothers bear a weapon</i>	SO	IA	2
2	Telum vulnerat fratres	<i>The weapon injures the brothers</i>	OS	IA	2
2	Victores ferrum portant	<i>The winners bear a sword</i>	SO	AI	3
2	Voctores ferrum vulnerat	<i>The sword injures the winners</i>	OS	AI	3
2	Telum fratres portant	<i>The brothers bear a weapon</i>	SO	IA	3
2	Telum fratres vulnerat	<i>The weapon injures the brothers</i>	OS	IA	3
3	Vident senes bellum	<i>The old men see the war</i>	SO	AI	1
3	Terret senes bellum	<i>The war scares the old men</i>	OS	AI	1
3	Vident periculum uxores	<i>The women see the danger</i>	SO	IA	1
3	Terret periculum uxores	<i>The danger scares the women</i>	OS	IA	1
3	Senes vident bellum	<i>The old men see the war</i>	SO	AI	2
3	Senes terret bellum	<i>The war scares the old men</i>	OS	AI	2
3	Periculum vident uxores	<i>The women see the danger</i>	SO	IA	2
3	Periculum terret uxores	<i>The danger scares the women</i>	OS	IA	2
3	Senes bellum vident	<i>The old men see the war</i>	SO	AI	3
3	Senes bellum terret	<i>The war scares the old men</i>	OS	AI	3
3	Periculum uxores vident	<i>The women see the danger</i>	SO	IA	3
3	Periculum terret uxores	<i>The danger scares the women</i>	OS	IA	3
4	Laudant senatores consilium	<i>The senators praise the resolution</i>	SO	AI	1
4	Iuvat senatores consilium	<i>The resolution pleases the senators</i>	OS	AI	1
4	Laudant templum patres	<i>The fathers praise the temple</i>	SO	IA	1
4	Iuvat templum patres	<i>The temple pleases the fathers</i>	OS	IA	1
4	Senatores laudant consilium	<i>The senators praise the resolution</i>	SO	AI	2
4	Senatores iuvat consilium	<i>The resolution pleases the senators</i>	OS	AI	2
4	Templum laudant patres	<i>The fathers praise the temple</i>	SO	IA	2
4	Templum iuvat patres	<i>The temple pleases the fathers</i>	OS	IA	2
4	Senatores consilium laudant	<i>The senators praise the resolution</i>	SO	AI	3
4	Senatores consilium iuvat	<i>The resolution pleases the senators</i>	OS	AI	3

Set	Example stimuli	Literal translation	Argument order	Animacy order	Verb position
4	Templum patres laudant	<i>The fathers praise the temple</i>	SO	IA	3
4	Templum patres iuvat	<i>The temple pleases the fathers</i>	OS	IA	3

References

- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. [https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)
- Bader, M., & Bayer, J. (2006). *Case and linking in language comprehension: Evidence from German: Vol. volume 34*. Springer Science & Business Media.
- 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. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates D, Mächler M, Bolker B, Walker S (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*. <http://arxiv.org/abs/1406.5823>
- Boeg Thomsen, D., & Poulsen, M. (2015). Cue conflicts in context: Interplay between morphosyntax and discourse context in Danish preschoolers' semantic role assignment. *Journal of Child Language*, 42(6), 1237–1266. <https://doi.org/10.1017/S0305000914000786>
- Brandt, S., Lieven, E., & Tomasello, M. (2016). German Children's use of word order and case marking to interpret simple and complex sentences: Testing differences between constructions and lexical items. *Language Learning and Development*, 12(2), 156–182. <https://doi.org/10.1080/15475441.2015.1052448>
- Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, 125, 104334. <https://doi.org/10.1016/j.jml.2022.104334>
- Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), 9. <https://doi.org/10.5334/joc.10>
- Chan, A., Lieven, E., & Tomasello, M. (2009). Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics*, 20, 267–300.
- Clahsen, H., & Felser, C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, 27(1), 107–126. <https://doi.org/10.1017/S0142716406060206>
- 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. <https://doi.org/10.1111/j.1467-8624.2008.01181.x>
- Dröge, A., Rabs, E., Fleischer, J., Billion, S. K. H., Meyer, M., Schmid, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2020). Case syncretism, Animacy, and word order in continental west Germanic: Neurolinguistic evidence from a comparative study on standard German, Zurich German, and Fering (North Frisian). *Journal of Germanic Linguistics*, 32(3), 217–310. <https://doi.org/10.1017/S1470542719000199>
- Dussias P E, Cramer T R (2006) The Role of L1 Verb Bias in L2 Sentence Parsing. In *Proceedings of the 30th annual Boston university conference on language development* (Vol. 1, pp. 166–177).
- Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27(2), 164–194. <https://doi.org/10.1093/applin/aml015>
- Ellis, N. C., & Sagarra, N. (2010). THE bounds of adult language acquisition: blocking and learned attention. *Studies in Second Language Acquisition*, 32(4), 553–580. <https://doi.org/10.1017/S0272263110000264>
- Ellis, N. C., & Sagarra, N. (2011). learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition*, 33(4), 589–624. <https://doi.org/10.1017/S0272263110000325>
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505. <https://doi.org/10.1111/j.1469-8986.2007.00531.x>
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (pp. 559–586). Lawrence Erlbaum Assoc.
- Frenc-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research*, 21(2), 175–198. <https://doi.org/10.1191/0267658305sr257oa>
- Frenc-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology A*, 50(1), 119–148. <https://doi.org/10.1080/027249897392251>
- Grünloh, T., Lieven, E., & Tomasello, M. (2011). German children use prosody to identify participant roles in transitive sentences. *Cognitive Linguistics*. <https://doi.org/10.1515/cogl.2011.015>

- Havik, E., Roberts, L., van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, 59(1), 73–112. <https://doi.org/10.1111/j.1467-9922.2009.00501.x>
- Hemforth, B., Konieczny, L., & Scheepers, C. (1993). *On reanalysis. Reanalysis and Repair*.
- Henry, N., Hopp, H., & Jackson, C. N. (2017). Cue additivity and adaptivity in predictive processing. *Language, Cognition and Neuroscience*, 32(10), 1229–1249. <https://doi.org/10.1080/23273798.2017.1327080>
- Henry, N., Jackson, C. N., & Hopp, H. (2020). Cue coalitions and additivity in predictive processing: The interaction between case and prosody in L2 German. *Second Language Research*. <https://doi.org/10.1177/0267658320963151>
- Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, 22(3), 369–397. <https://doi.org/10.1191/0267658306sr272oa>
- Hopp, H. (2013). The development of L2 morphology. *Second Language Research*, 29(1), 3–6. <https://doi.org/10.1177/0267658312465304>
- Hopp, H. (2015a). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306.
- Hopp, H. (2015b). Individual differences in the second language processing of object–subject ambiguities. *Applied Psycholinguistics*, 36(2), 129–173. <https://doi.org/10.1017/S0142716413000180>
- Hopp, H. (2020). Morphosyntactic adaptation in adult L2 processing: Exposure and the processing of case and tense violations. *Applied Psycholinguistics*, 41(3), 627–656. <https://doi.org/10.1017/S0142716420000119>
- Jackson, C. (2008). Proficiency Level and the interaction of lexical and morphosyntactic information during L2 sentence processing. *Language Learning*, 58(4), 875–909. <https://doi.org/10.1111/j.1467-9922.2008.00481.x>
- Jackson, C. N., & Dussias, P. E. (2009). Cross-linguistic differences and their impact on L2 sentence processing. *Bilingualism*, 12(1), 65–82. <https://doi.org/10.1017/S1366728908003908>
- Jackson, C. N., & Roberts, L. (2010). Animacy affects the processing of subject–object ambiguities in the second language: Evidence from self-paced reading with German second language learners of Dutch. *Applied Psycholinguistics*, 31(4), 671–691. <https://doi.org/10.1017/S0142716410000196>
- Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in wh-extraction. *Studies in Second Language Acquisition*, 17(4), 483–516. <https://doi.org/10.1017/S027226310001442X>
- Kaan, E. (2014). Predictive sentence processing in L2 and L1. *Linguistic Approaches to Bilingualism*, 4(2), 257–282.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156. [https://doi.org/10.1016/S0749-596X\(03\)00023-8](https://doi.org/10.1016/S0749-596X(03)00023-8)
- Keip, M., & Doepner, T. (2019). *Interaktive Fachdidaktik Latein (4)*. Vandenhoeck & Ruprecht.
- Kim, A., & Sikos, L. (2011). Conflict and surrender during sentence processing: An ERP study of syntax-semantics interaction. *Brain and Language*, 118(1–2), 15–22. <https://doi.org/10.1016/j.bandl.2011.03.002>
- Kimball, J. (1975). Predictive analysis and over-the-top parsing. *Syntax and Semantics*, 4, 155–179.
- Köhne, J., & Crocker, M. W. (2010). Sentence Processing Mechanisms Influence Cross-Situational Word Learning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 32(32), 2458–2463.
- Kuhlmann, P. (2015). Lateinische Texte richtig übersetzen—(K)ein Problem? Die lernpsychologischen Voraussetzungen für das Verstehen von lateinischen Texten. In M. Frisch (Ed.), *Alte Sprachen—Neuer Unterricht pp 11–35*. Kartoffeldruck-Verlag.
- Kuhlmann, P. (2019). Textverstehen und Übersetzen sowie deren Bedeutung. In T. Choitz, K. Sundermann, & A. Meyer (Eds.), *Perspektiven für den Lateinunterricht III* (pp. 38–49). Klett.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703.
- MacWhinney, B. (1992). Transfer and Competition in Second Language Learning. *Advances in Psychology* (pp. 371–390). Amsterdam: Elsevier.
- MacWhinney, B. (2002). The Competition Model: The Input, the Context, and the Brain. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.
- MacWhinney, B. (2005). A Unified Model of Language Acquisition. In J. F. Kroll & M. A. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford University Press.
- MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23(2), 127–150.
- MacWhinney, B., Leinbach, J., Taraban, R., & McDonald, J. (1989). Language learning: Cues or rules? *Journal of Memory and Language*, 28, 255–277.
- Marslen-Wilson, W. D. (1975). Sentence perception as an interactive parallel process. *Science*, 189(4198), 226–227.
- McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283–312.
- Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism*, 19(1), 19–35. <https://doi.org/10.1017/S1366728914000881>
- Myers, J. (2017). Acceptability judgements. *Oxford Research Encyclopedias Linguistics*. <https://doi.org/10.1093/acrefore/9780199384655.013.333>
- Papadopoulou, D. (2005). Reading-time studies of second language ambiguity resolution. *Second Language Research*, 21(2), 98–120. <https://doi.org/10.1191/0267658305sr254oa>
- Potter, M. C. (1984). Rapid Serial Visual Presentation (RSVP). A Method for Studying Language Processing. In D. E. Kieras & M. A. Just (Eds.), *New Methods in Reading*

- Comprehension Research* (pp. 91–118). England: Routledge.
- Sanz, C., Park, H. I., & Lado, B. (2015). A functional approach to cross-linguistic influence in *ab initio* L3 acquisition. *Bilingualism*, 18(2), 236–251. <https://doi.org/10.1017/S1366728914000285>
- Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (pp. 65–93). Kluwer.
- Stafford, C. A., Bowden, H. W., & Sanz, C. (2012). Optimizing language instruction: Matters of explicitness, practice, and cue learning: instructional explicitness and cue learning. *Language Learning*, 62(3), 741–768. <https://doi.org/10.1111/j.1467-9922.2011.00648.x>
- Staub, A., & Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from either or. *Journal of Experimental Psychology*, 32(2), 425–436. <https://doi.org/10.1037/0278-7393.32.2.425>
- Neurobehavioral Systems. (2016). *Presentation software, version 19.0*. www.neurobs.com
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a Lexicalist Framework for Constrain-Based Syntactic Ambiguity Resolution. *Perspectives on Sentence Processing* (pp. 155–180). Lawrence Erlbaum Assoc.
- VanPatten, B., & Smith, M. (2019). Word-order typology and the acquisition of case marking: A self-paced reading study in Latin as a second language. *Second Language Research*, 35(3), 397–420. <https://doi.org/10.1177/0267658318785652>
- Weiss, A. F. (2020). The information gathering framework—A cognitive model of regressive eye movements during reading. *Journal of Eye Movement Research*. <https://doi.org/10.16910/jemr.13.4.4>
- Williams, J. N., Möbius, P., & Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics*, 22, 509–540.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.