Research Article

Peter de Swart*, Geertje van Bergen

How Animacy and Verbal Information Influence V2 Sentence Processing: Evidence from Eye Movements

Abstract: There exists a clear association between animacy and the grammatical function of transitive subject. The grammar of some languages require the transitive subject to be high in animacy, or at least higher than the object. A similar animacy preference has been observed in processing studies in languages without such a categorical animacy effect. This animacy preference has been mainly established in structures in which either one or both arguments are provided before the verb. Our goal was to establish (i) whether this preference can already be observed before any argument is provided, and (ii) whether this preference is mediated by verbal information. To this end we exploited the V2 property of Dutch which allows the verb to precede its arguments. Using a visual-world eye-tracking paradigm we presented participants with V2 structures with either an auxiliary (e.g. *Gisteren heeft X ...* 'Yesterday, X has ...') or a lexical main verb (e.g. *Gisteren motiveerde X ...* 'Yesterday, X motivated ...') and we measured looks to the animate referent. The results indicate that the animacy preference can already be observed before arguments are presented and that the selectional restrictions of the verb mediate this bias, but do not override it completely.

Keywords: Verb-initial sentences; predictive processing; visual world; Dutch; lexical information; selectional restrictions; argument structure; grammar-performance connection

1 Introduction

What we talk about influences how we talk about it. This especially is true of the properties of the referents we talk about. Animacy is one recurrent referential feature that influences the form of linguistic expressions. Animacy concerns the property of being alive or sentient and is known to affect a wide variety of linguistic domains, including grammatical gender (Dahl 2000, Siemund 2008), agreement (Ortmann 1998), case (de Swart 2007, Malchukov 2008, Fauconnier 2011) and word order (Tomlin 1986, Siewierska 1988, Branigan et al. 2008, van Bergen 2011). Thus, we find that many languages of different genealogical and geographic background only overtly mark direct objects that refer to animate entities, but not those that refer to inanimate entities (*differential object marking*, Aissen 2003, de Swart 2007, Malchukov 2008). This is just one example of many. In fact, the effect of animacy seems "so pervasive in the grammars of human languages that it tends to be taken for granted and become invisible" (Dahl & Fraurud 1996: 47).

The concept of animacy and hence the distinction between animate and inanimate entities is often hard to define and classification may change from one language or even situation to the next (e.g. Yamamoto 1999). Nevertheless, linguists quite consistently speak about animacy proper in terms of a three-way

*Corresponding author: Peter de Swart, Centre for Language Studies, Radboud University Nijmegen, Nijmegen, 6500 HD, The Netherlands, E-mail: p.deswart@let.ru.nl
Geertje van Bergen, Max Planck Institute for Psycholinguistics, Nijmegen, 6525 XD, The Netherlands

hierarchy ranking humans above animals (non-human animates), which in turn are ranked above inanimate entities: human > animate > inanimate (cf. Comrie 1989). Discussion exists about the grammatical nature and universality of such prominence hierarchies (see Lockwood & Macaulay 2012, and the contributions to Bornkessel-Schlesewsky et al. 2015).

We will focus in this paper on what is known as the relational use of the animacy hierarchy: in many languages two arguments have to be compared with respect to the hierarchy in order to determine the grammatically licensed form or interpretation of an utterance. A classic example that falls into this category is word order in Navajo (e.g. Young & Morgan 1987). In this language the element that ranks highest on the animacy hierarchy has to occur first in the sentence. A related restriction is reported for Mam-Maya (Minkoff 2000). In this language the subject of an active transitive sentence has to rank highest on the hierarchy. This means that if the agent ranks lower on the hierarchy than the patient, a speaker has to revert to a passive construction (e.g. *the man was seen by the dog*). These two examples seem to fit into a wider cross-linguistic pattern in which the assignment of Agent (subject) and Patient (object) is determined by the generalization in (1), which we will refer to as the *animacy preference*:

(1) Agent should outrank Patient on the animacy hierarchy H_l
(where H_l is a language-specific ranking)

The association of entities higher in animacy with the Agent role is not surprising, given that several agentive properties, such as volitionality and sentience, entail animacy (see Primus 2012 for discussion).

We will be concerned with the reflection of the generalization in (1) in languages in which the animacy hierarchy has not grammaticalized, or more precisely, does not have a categorical effect, in contrast to Navajo and Mam-Maya and many other languages. It has long been observed that animacy induces statistical tendencies in language use that are very similar to the categorical effects discussed for the languages above. Bresnan et al. (2001) dubbed this pattern 'hard constraints mirror soft constraints' and took it as a motivation for a grammatical model in which constraints can be ranked lower, resulting in statistical patterns, or high, resulting in categorical patterns. This approach is reminiscent of the notion of cue validity in the Competition model (MacWhinney & Bates 1989). The idea that there exists a connection between cross-linguistic grammatical patterns and patterns in language use is, for instance, also incorporated in the Performance-Grammar Correspondence Hypothesis of Hawkins (2004) and the Interface Hypothesis of Incremental Argument Interpretation of Bornkessel-Schlesewsky & Schlesewsky (2009).

The grammatical patterns in Mam-Maya and Navajo were, for instance, found to be mirrored in sentence production in Germanic languages and beyond, as witnessed by corpus frequencies and controlled experiments (e.g., Øvrelid 2004, Kempen & Harbusch 2004, Snider & Zaenen 2006, see also Branigan et al. 2008 for discussion). The effect of animacy on language comprehension is also well-documented. Animacy is, for instance, known to influence the processing difficulty associated with object-relative clauses in comparison to subject-relative clauses (Trueswell et al 1994, Mak et al. 2002, Mak et al. 2006). Both Mak et al. (2006) for Dutch and Wu et al. (2012) for Mandarin Chinese propose an animacy preference constraint that prefers animate subjects over inanimate ones. Similar effects have been observed for simple transitive sentences. Bornkessel-Schlesewsky & Schlesewsky (2009) provide an overview of cross-linguistic investigations showing the relational effect of animacy on sentence processing. In particular, when an animate P-argument has been unambiguously identified and is being followed by an inanimate A-argument this results in additional processing costs. They incorporated a constraint as in (1) above into their model to guide role assignment during on-line language processing. Finally, Paczynski & Kuperberg (2011) propose that a direct mapping from the animacy hierarchy to linear order affects processing, such that less animate arguments should appear later (the animate-first effect). They observed increased processing costs for English sentences with an animate direct object in comparison to those with an inanimate one. The effect was independent of semantic role assignment.

The influence of animacy on language comprehension in languages without categorical animacy effects is thus firmly established. Most evidence for the reliance on the animacy preference in (1) in language use comes from studies based on verb-final or verb-medial structures. This means that the observed effect

of animacy is completely or partially independent from verbal information. Bornkessel-Schlesewsky & Schlesewsky (2009) argue that this verb-independent nature is a prerequisite for efficient communication. Given that many languages have a verb-final word order, the use of an independent heuristic such as that in (1) above, allows language users to establish an interpretation without delay and before the verb has been encountered. Nevertheless, the verb may have a profound effect on the assignment of semantic roles through its selectional restrictions. Due to close connections between animacy and certain role types, such as agent and experiencer (e.g. Van Valin & LaPolla 1997, Primus 2012), animacy characteristics in combination with the argument structure of the verb may guide the interpretation of sentences. Thus, when an animate and inanimate argument are combined with an agentive verb like *break* the former will be associated with the Agent slot and hence the subject function in an active sentence, but when combined with an object-experiencer verb like *please* it will be associated with the object function. This suggests that the preference in (1) can be counteracted by the verbal argument structure (e.g. Lamers 2012, Lamers & de Hoop 2014, Czypionka 2014). The main goal of the current study is to investigate this interaction between verbal information and the animacy preference in (1). In addition, we want to establish whether the animacy preference in (1) is already observed before any of the arguments have been encountered. The studies discussed above inferred the significance of the animacy preference for language comprehension from the effects that emerge when one or both of the arguments have been processed. The significance of the animacy preference will be strengthened when similar effects can be observed before any arguments have been encountered. If language users indeed rely on animacy information to establish an (expectation of an) interpretation, we would expect to find this reflected in measures of predictive processing. This would provide complementary evidence that the animacy preference in (1) is used as a heuristic during sentence comprehension.

It is well known by now that online language comprehension is not a passive integration process: readers and listeners use various types of (non)linguistic cues to predict upcoming input (for reviews, see e.g., Federmeier 2007, Kamide 2008, van Petten & Luka 2012, Pickering & Garrod 2013, Kuperberg & Jaeger 2016, among many others). Ways of measuring predictive processing include anticipatory eye movements to depicted referents while listening to speech (Visual World eye-tracking, e.g., Tanenhaus et al. 1995, Altmann & Kamide 1999, see also Salverda & Tanenhaus 2017), and modulations of event-related potentials (ERPs) on linguistic elements preceding the critical word during reading/listening (e.g., DeLong et al. 2005, van Berkum et al. 2005). For instance, in a seminal study using Visual World eye-tracking, Altmann & Kamide (1999) showed that language comprehenders use the selectional restrictions of the verb as a cue to predict which direct object will follow. While listening to sentences like *The boy will eat ...*, participants fixated the only edible object (cake) in the visual display more often than non-edible referents, well before they heard the target word. These semantic restrictions of the verb can also be animacy-based, e.g., *The policeman will arrest ...* can only be followed by an animate argument. Other evidence for animacy-based predictions during incremental language comprehension comes from Szewczyk & Schriefers (2013). Using an ERP reading study, they showed that native Polish listeners use animacy information to predict upcoming direct objects following a subject and a transitive verb. In their experiment, participants read contexts that were biasing towards either an animate or an inanimate direct object. If the animacy of the direct object did not match the expected animacy value, they found enhanced neural activity (N400) on the (differentially inflected) adjective preceding the critical noun, that is, before the direct object had been fully encountered.

Whereas the above studies demonstrate that animacy information is indeed used to anticipate upcoming arguments, note that the prediction effects were established after the subject and the verb had already been processed. As a result, the animacy information was no longer needed for argument disambiguation. Sauppe (2016) correctly points out that subject-initial (i.e., verb-medial or verb-final) languages conflate different information types (i.e., grammatical function, semantic role, word order) on a single argument, i.e., the direct object, hence concealing the precise nature of prediction processes. In order to disentangle effects of syntactic function, word order and thematic role, he focuses on the verb-initial language Tagalog. Using a Visual World eye-tracking paradigm similar to Altmann & Kamide (1999), he found that anticipatory looks were influenced by the semantic rather than the syntactic information from the verb: after having heard the verb, listeners fixated the agent irrespective of its syntactic function or linear position in the sentence.

Based on his findings, he stresses the importance of the prominence of the agent role in processing and predicting event structures (e.g., Kemmerer 2012, Cohn & Paczynski 2013, Bornkessel-Schlesewsky & Schlesewsky 2013, cf. also Primus 2012), and argues that an investigation of more diverse languages in the world will help to refine psycholinguistic theories.

We underline Sauppe's (2016) plea to make use of typological diversity in language research in general, but in this case very similar and even more detailed insights can be deduced from 'usual suspect' languages (Levinson 2012, Norcliffe et al. 2015), i.e., verb-final languages like Dutch and German, by exploiting their verb second (V2) property. The V2 property dictates that the finite verb takes up the second position in main clauses. The finite verb can either be an auxiliary (as in (2a)) or a lexical finite verb (2b,c) (the finite verb is underlined):

(2) a.   *Zoals verwacht __**heeft**__ de medaille de zwemmer **gemotiveerd**.*
           as expected has the medal the swimmer motivate.PTCP
           'As expected, the medal motivated the swimmer.'

   b.   *Zoals verwacht __**motiveerde**__ de medaille de zwemmer.*
           as expected motivate.PST the medal the swimmer
           'As expected, the medal motivated the swimmer.'

   c.   *Zoals verwacht __**ontving**__ de zwemmer de medaille.*
           as expected receive.PST the swimmer the medal
           'As expected, the swimmer received the medal.'

Dutch grammar dictates that the first NP following the finite verb has to be interpreted as the subject (unlike German which allows subject-object reversal, cf. Lenerz 1977). This means that upon encountering the finite verb, Dutch listeners will anticipate a subject. Crucially, however, the difference in lexical information between the two configurations in (2) will affect predictions about which argument will be the subject. The auxiliary in (2a) will be neutral, or at most elicit a preference for an agentive or (through the connection between agentivity and animacy) animate subject, but the lexical finite verb in (2b) allows for more fine-grained preferences. Given the animate and inanimate arguments in (2b), the only possible configuration upon hearing the verb is one where the swimmer takes up the direct object position and hence will be named second. Changing the verb to *ontvangen* ('to receive), as in (2c), will anticipate a reverse interpretation, assigning the subject function to the swimmer. The V2 property thus allows us to compare the effect of verbal information in verb-initial contexts with that in verb-final contexts within one language. This provides a clean view on the precise role of verbal information and its interaction with the animacy preference.

Surprisingly little work has been reported on the effect of verbal information in the (Germanic) V2 position. Scheepers et al. (2000) investigated the effect of verb position on the comprehension of German SO and OS sentences with different types of verbs. The position of the verb did not have an effect on the acceptability of the sentences (Experiment 1): in both cases, SO sentences were rated higher than OS sentences. Moreover, it did not affect the reading times in an eye-tracking-during-reading experiment (Experiment 3). Bader & Bayer (2006, Chapter 6, Experiment 2), by contrast, report an effect of verb position across experiments. They found that globally ambiguous sentences with the lexical verb in second position elicited a garden-path effect when disambiguated towards both an SO reading and an OS reading. However, when the second position was taken up by an auxiliary (Chapter 6, Experiment 1) this only elicited a garden-path effect for OS sentences. Moreover, the garden-path effect was weaker for sentences with the lexical verb in second position. This suggests that processing is affected by verbal information preceding two arguments, attenuating the generally strong SO preference found in German.

Dahan & Tanenhaus (2004) also provide evidence for the use of verbal information in the V2 position, even though this was not their main objective. Using a Visual World experiment, they targeted effects of the selectional restrictions of verbs on word recognition in Dutch. They used intransitive sentences with either the neutral auxiliary *is* 'is' (the non-constraining context) or a lexical verb (e.g. *klom* 'climbed', the constraining context) in second position (V2), and compared fixations to a target referent (e.g. *bok* 'male goat') and a phonological competitor (*bot* 'bone'). Fixations to the target and the phonological competitor

were found to diverge much earlier in the constraining context compared to the non-constraining context, which strongly suggests that listeners readily use thematic constraints derived from verbal information ('only a goat can climb').

Both the findings of Bader & Bayer (2006) and those of Dahan & Tanenhaus (2004) suggest that verbal information in the second sentence position is used by the language comprehension system. Since Dahan and Tanenhaus focused on intransitive sentences, and Bader and Bayer considered only sentences with two animate arguments, it remains as of yet unclear how the animacy preference in (1) interacts with verbal information in the incremental understanding of transitive sentences. In the current study, we address this question by presenting Dutch listeners with transitive sentences as the ones in (2) above, using a methodology similar to Dahan & Tanenhaus (2004). We will use sentences containing an animate and an inanimate argument that are either the subject or object, depending on the verb (e.g. *ontvangen* 'receive' vs. *motiveren* 'motivate'); the lexical verb will take up either the V2 position or the verb final position. These manipulations result in two predictions. First, in the absence of lexical verbal information we expect more fixations to the animate referent than to the inanimate referent. This would reflect the animacy preference, and the preference for an agentive reading associated with the auxiliary *heeft* (as explained above). Second, if listeners immediately use the verbal information in the verb to predict upcoming arguments, we hypothesize that the animacy preference is less strong in the case of verbs that select an inanimate subject compared to verbs that select an animate subject, which would pattern similar to the auxiliary verb.

2 Method

2.1 Participants

Eighty-nine students of the University of Groningen participated in the experiment (73 females, 16 males, mean age = 21.8 years); two participants had to be excluded because of technical error. All were native speakers of Dutch. They were paid 5 EUR for their participation.

2.2 Materials and design

2.2.1 Experimental stimuli

We constructed 16 sets of experimental items, consisting of a combination of an animate noun - inanimate noun pair with two verbs, one taking an animate subject and an inanimate object (*animate* verb), the other one taking the opposite pattern (*inanimate* verb). Although many verbs could in principle take both animate and inanimate nouns as either their subject or object, the specific combination with the selected noun pairs left only one plausible configuration per verb (see Appendix A for a list of all noun-verb combinations used). For each item set we constructed four sentences, manipulated for ANIMACY CONFIGURATION (*animate subject – inanimate object* vs. *inanimate subject – animate object*) and VERB TYPE (*auxiliary* vs. *lexical*), referring to the finite verb in V2 position. All sentences followed the template in (3):

> (3) XP - Vfin - NP1[=subject] - NP2[=object] - (Past Participle) - PP

In sentences containing an auxiliary, the lexical verb occurred as a past participle directly following NP2; in sentences with the lexical verb in V2 position, the past participle was omitted. An example of an experimental item set in each condition is given in (4):

> (4) a.   AUXILIARY FINITE VERB * ANIMATE SUBJECT – INANIMATE OBJECT
>   *Zoals verwacht heeft de zwemmer de medaille ontvangen na het toernooi.*
>   as expected has the swimmer the medal receive.PTCP after the tournament

          'As expected, the swimmer received the medal after the tournament.'

b.   AUXILIARY FINITE VERB * INANIMATE SUBJECT – ANIMATE OBJECT
          Zoals verwacht heeft de medaille de zwemmer gemotiveerd na het toernooi.
          as expected has the medal the swimmer motivate.PTCP after the tournament
          'As expected, the medal motivated the swimmer after the tournament.'

c.   LEXICAL FINITE VERB * ANIMATE SUBJECT – INANIMATE OBJECT
          Zoals verwacht ontving de zwemmer de medaille na het toernooi.
          as expected receive.PST the swimmer the medal after the tournament
          'As expected, the swimmer received the medal after the tournament.'

d.   LEXICAL FINITE VERB * INANIMATE SUBJECT – ANIMATE OBJECT
          Zoals verwacht motiveerde de medaille de zwemmer na het toernooi.
          as expected motivate.PST the medal the swimmer after the tournament
          'As expected, the medal motivated the swimmer after the tournament.'

The plausibility of verb-noun pair combinations was assessed in a pretest. Thirty participants who did not participate in the actual experiment rated the plausibility of the sentences in the present perfect form (cf. 4a and 4b) on a 7-point scale (7=highly plausible). A linear mixed-effect analysis with maximal random effect structure revealed no difference ($p = 0.15$) between the two types of ANIMACY CONFIGURATION (animate subject-inanimate object $M = 5.64$, $SD = 1.44$; inanimate subject-animate object $M = 5.41$, $SD = 1.78$).

Experimental items were counterbalanced across four lists following a Latin square design, such that each participant heard each verb-noun pair combination in only one condition. VERB TYPE was manipulated between subjects, i.e., half of the participants heard only sentences with the auxiliary in V2 position ((4a) and (4b)) and the other half was exposed only to sentences with the lexical verb in V2 position ((4c) and (4d)). ANIMACY CONFIGURATION was varied within subjects, such that each participant encountered each animacy configuration eight times. Each list occurred in two pseudo-randomized orders.

2.2.2 Filler items

In addition to the experimental items we constructed 40 filler sentences, which were constructed using the general format in (3) above. Eight items included an animate referent acting on another animate referent (e.g. *The mailman greeted the butcher*); eight items presented two inanimate referents (e.g. *the towel covered the bucket*). Twenty-four items involved animals and/or vehicles acting on one another, all of which were part of another experiment. Twelve of these items consisted of animals and vehicles pulling or pushing one another (e.g. *the cow pushed the bus*, *the motorboat pulled the seal*), six items involved two animals acting on another (e.g. *the cat grabbed the pigeon*), and six items figured two vehicles (e.g. *the police car overtook the truck*). Six practice trials were constructed representing a combination of experimental and filler items.

2.2.3 Visual stimuli

Visual displays contained two images depicting the referents of the nouns (see Bosker et al. 2014 for a similar set-up). We opted for two-picture displays without any competitors, as competitors would also have an animacy value and hence would introduce an effect of plausibility. Photographs were acquired through Google Image searches and from the photo-stock website Dreamstime.com, and edited (e.g., background removed where possible) to reduce visual complexity; all were converted to grey-scale and resized to a square format (675x675 pixels) to minimize effects of visual salience. To control for looking order, the position of the pictures on the screen (animate left vs. animate right) was counterbalanced.

2.2.4 Auditory stimuli

Sentences were recorded by a male native speaker of Dutch (age = 29 years), speaking at normal speed. To reduce the dynamic range of the audio recordings the audio files were leveled using the Audacity software. To increase the time between the offset of the verb and the onset of the following NP, a silent pause (duration 130 ms) was manually added to the sound file. For the auxiliary conditions, the same instance of the verb (*heeft)* including the 130 ms pause was spliced into all audio recordings.

2.3 Procedure

The experiment was conducted in the EyeLab at the University of Groningen. Eye-movements of both eyes were recorded with a Tobii T120 eye tracker sampling at 120Hz. The experiment was programmed in E-prime2 with the extensions for Tobii (Psychology Software Tools 2009). Participants were seated in a semi-dark room at approximately 65cm from the screen, wearing headphones for the audio stimuli.

After a short calibration phase, the participants were first familiarized with the pictures and the names of their referents. They saw each picture and heard the corresponding name spoken by the same male native speaker of Dutch who recorded the sentence stimuli. The familiarization task was self-paced, as participants could proceed by hitting the space bar. Success of familiarization was not recorded.

Experimental trials (Figure 1) started with a gaze-contingent fixation cross presented at the center of the screen. When participants fixated the cross, it disappeared and the two pictures of the referents appeared on the screen at equal horizontal distance from the center. After a preview window of 300ms, the sound file with the corresponding stimulus sentence started to play. Participants were instructed to look at the screen and listen carefully to the sentences. In order to assure that participants were paying attention to the audio stream, we included questions on 37 of the 61 trials. These questions were statements about the sentences, which could be either true or false. Participants had to respond by pressing marked keys on the keyboard ('c' for correct and 'm' for false). Incorrect statements were formed by replacing either one of the nouns, the verb, the adverbial expression or the prepositional phrase, or by reversing the direction of the action. The experiment lasted approximately 30 minutes.
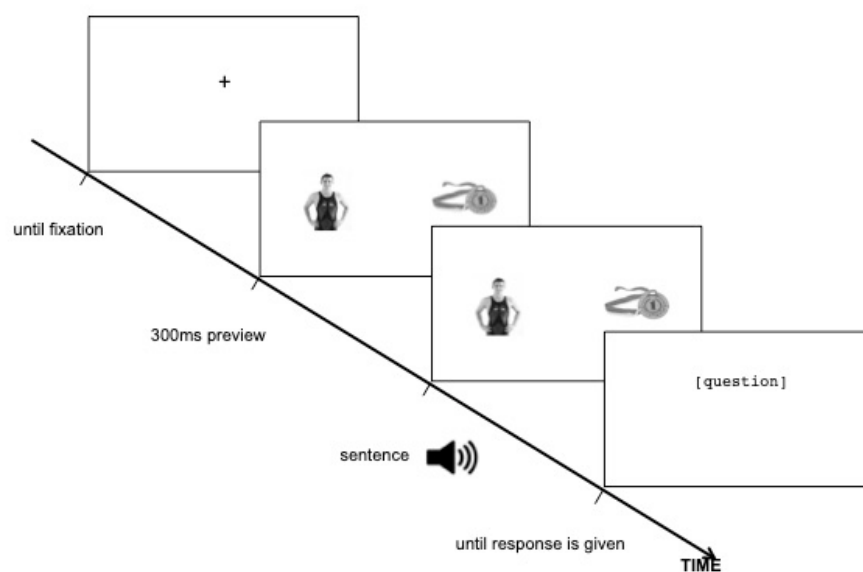

Figure 1: The structure of a trial.

2.4 Analysis

We manually determined the onsets and offsets of the finite verb, as well as the onsets of NP1 and NP2, using the Audacity and Praat software (Broersma & Weenink 2016). This information was used to time-lock fixation proportions to different time points in the sentence. Fixations of less than 50ms were excluded, as were fixations outside the areas of interest. We analyzed fixations in four time windows (see Figure 2): (i) the pre-verbal window, (ii) the pre-argument window, (iii) the subject integration window, and (iv) the object integration window.

Fixations in the *pre-verbal window* were analyzed to establish initial looking preferences, without any influence of the following verb or arguments (before the verb, sentences are identical in all experimental conditions). This window runs from sentence onset until 200ms after the onset of the verb (mean duration 1125ms). As it takes minimally 200ms to plan and launch an eye movement (Salverda et al. 2014), we take 200ms after verb onset to be the earliest point in time at which fixations could in principle be affected by verbal information.

We analyzed gaze behavior in the *pre-argument* window to assess anticipatory eye movements towards upcoming subject arguments. This window runs from verb offset until 200ms after subject onset (duration 330ms). This window is most crucial to our research question as it will speak directly to the use of verbal information and its interaction with the animacy preference. We did not choose verb onset + 200ms as the starting point of this time window because of the differences in length between the auxiliary (duration 310ms) and the lexical verbs (mean duration 704ms): after 310ms, the verb would have been fully encountered in the auxiliary condition, but not in the lexical verb condition. Verb offset is the moment at which we can be sure that the verb has been completely encountered in both conditions. Although the earliest changes in fixations can only be detected after 200ms, the influence of verbal information on gaze behavior probably starts before the verb has been completely perceived, given that many, and especially multisyllabic words (which 27 out of 32 lexical verbs were) are recognized well before word offset (Marslen-Wilson, 1987). Moreover, the verbs occurred in the past tense, which for regular verbs (75% of the lexical
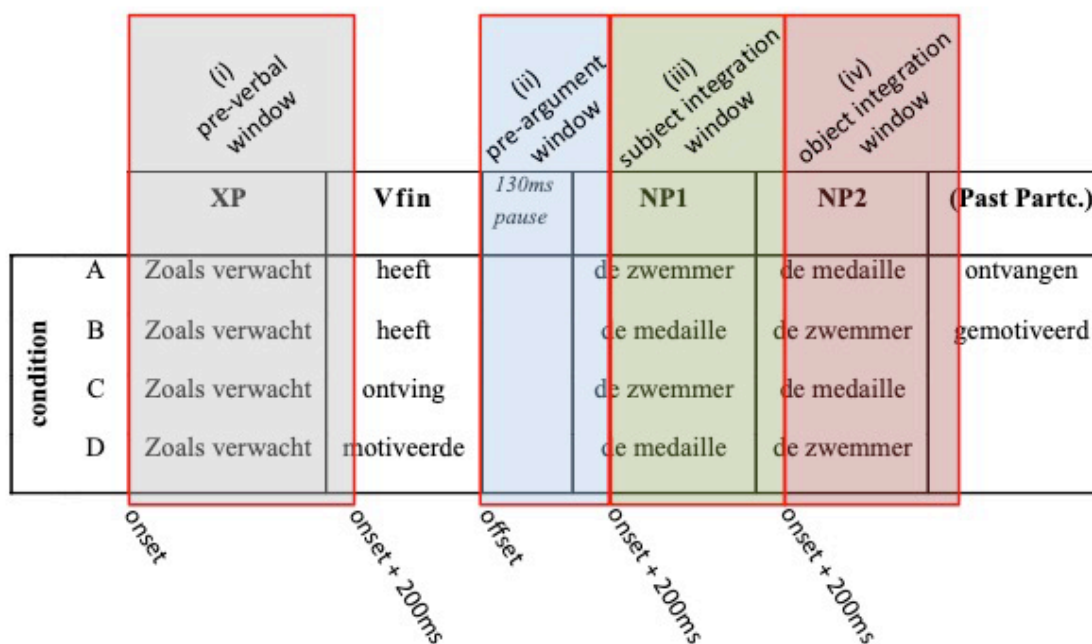
		XP	V fin	130ms pause	NP1	NP2	(Past Partc.)
condition	A	Zoals verwacht	heeft		de zwemmer	de medaille	ontvangen
	B	Zoals verwacht	heeft		de medaille	de zwemmer	gemotiveerd
	C	Zoals verwacht	ontving		de zwemmer	de medaille	
	D	Zoals verwacht	motiveerde		de medaille	de zwemmer	

Figure 2: Schematic representation of the windows of analysis with an example sentence per condition: (i) the pre-verbal window (sentence onset – verb onset + 200ms), (ii) the pre-argument window (verb offset – NP1 onset + 200ms), (iii) the subject integration window (NP1 onset + 200ms – NP2 onset + 200ms), and (iv) the object integration window (NP2 onset – duration of NP2).

verbs) means that they end in an inflectional suffix (-*de/-te*), which does not contribute to the lexical content of the verb. The verb in the auxiliary condition (*heeft*) is relatively short, but the same verb occurred in all sentences; as a result of this repetition, *heeft* was probably also recognized well before its offset. Taken together, this time window provides the least biased window across conditions to detect changes in gaze patterns related to verbal information.

Additionally, we analyzed fixations in the two argument windows, to explore whether and when differences between conditions in the pre-argument window affected the integration of subsequent arguments. As our study was not designed to test the effects on integration directly, the analyses of these windows should be considered exploratory. Nevertheless, this exploration may provide initial insight on how the availability of verbal information affects subsequent processing. Clearly, future studies have to be designed that target this issue directly.

The *subject integration window* was divided into two equal parts, i.e., an early integration window and a late integration window. These windows were determined by calculating for each trial the duration from NP1 onset + 200ms until NP2 onset + 200ms, and dividing it into two equal halves. An early and a late *object integration window* were created with a similar logic, starting at NP2 onset + 200ms. Since the noun pairs we used were identical across conditions, we obtained NP2 durations by taking the duration of the corresponding NP1. Hence, both argument integration windows used the same duration per item (M = 640ms, SD = 90.8).

For each time window, we calculated the total duration of AOI fixations per trial and per participant, and computed the logit-transformed odds ratio of fixations to the animate referent over those to the inanimate referent. We used a linear mixed effects regression model to predict the probability of fixations to the animate referent over those to the inanimate referent based on the fixed factors ANIMACY CONFIGURATION and VERB TYPE and their interaction. Picture order (animate left, animate right) and trial order were included as control variables. Factors were centered or sum coded. Our final models included the maximal random effects structure justified by the data (Barr et al. 2013).

3 Results

One participant was discarded from the analysis because they did not reach the threshold of 80% correct on target item questions. For the remainder of the participants, mean accuracy on target items was very high (M = 95.1%, SD = 4.9). No further trials were excluded based on correctness of answering the question. Forty-three trials (3%) were excluded due to extensive track loss, defined as track loss for both eyes on more than one third of the time during the trial.

Figure 3 shows the fixation proportions to the animate over the inanimate referent in the four conditions, time-locked to subject onset.

3.1 Pre-verbal window

To assess whether there was an inherent animacy preference, we first analyzed the probability of fixating animate over inanimate referents. We did so by inspecting the intercept of the statistical model. As the intercept presents the grand mean, an intercept that equals zero would indicate that there is no inherent looking preference. The intercept did, however, deviate significantly from zero (β = 1.74, *SE* = 0.23, *p* < .001), indicating a preference to look at the animate referent. As expected, there was no significant effect of ANIMACY CONFIGURATION (*p* = .61) or VERB TYPE (*p* = .17); nor their interaction (*p* = .36).
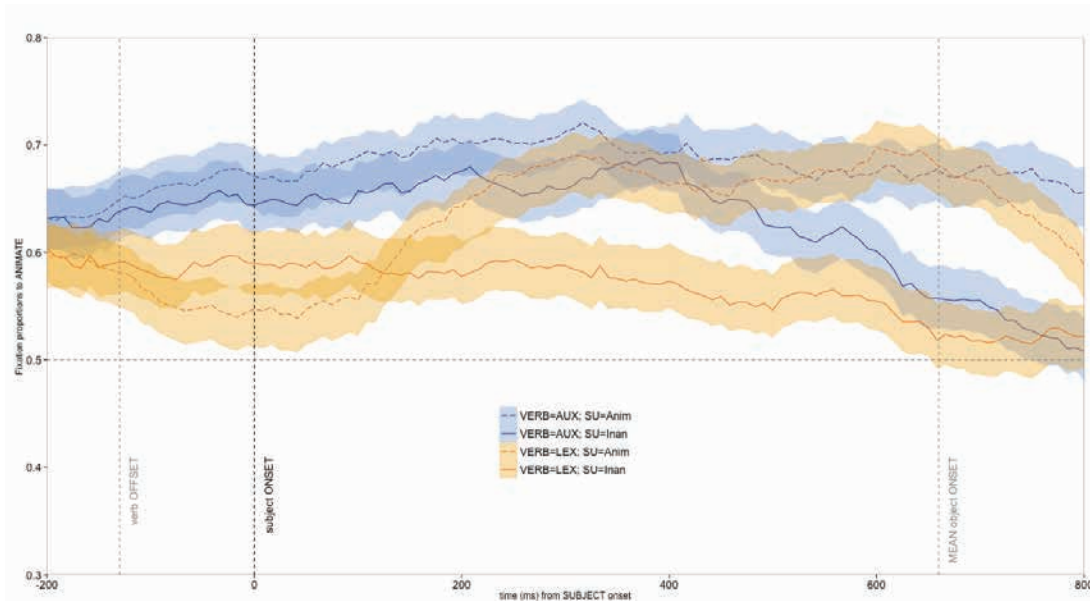
Figure 3: Gaze plot of animate fixation proportions in each condition time-locked to subject onset. The lines indicate mean fixation proportions and the transparent bands indicate 1 SE above/below the means. Blue lines indicate conditions with an auxiliary finite verb, orange lines indicate conditions with a lexical finite verb, dashed lines indicate conditions with an animate subject, and straight lines indicate conditions with an inanimate subject.

3.2 Pre-argument window

As Figure 3 shows, the animacy preference observed in the pre-verbal window seems to extend into this later window. Indeed, also in this case the intercept deviated significantly from zero ($\beta = 1.75$, $SE = .45$, $p = < .001$). In addition, the proportion of animate fixations seems to be higher for sentences containing an auxiliary in comparison to those with a lexical verb; this is also reflected in overall gaze durations (Figure 4). The statistical analysis confirmed that the main effect of VERB TYPE was significant ($\beta = 0.99$, $SE = 0.34$, $p = .0053$); we found no significant main effect of ANIMACY CONFIGURATION ($p = .62$), nor an interaction ($p = .61$).
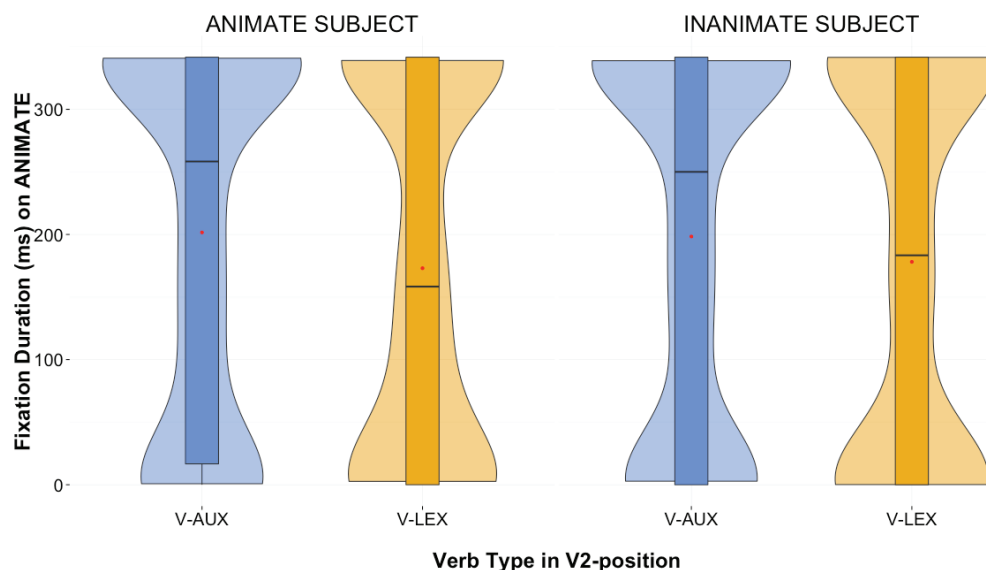

Figure 4: Violin/boxplot of fixation duration to the animate referent per condition in the pre-argument window. The red dot encodes the mean, the black vertical line encodes the median, the top and bottom edges of the box encode the interquartile range and the transparent beans provide a density plot. Note that the statistical analysis was performed on the logit-transformed odds ratio of fixations to the animate referent over those to the inanimate referent.
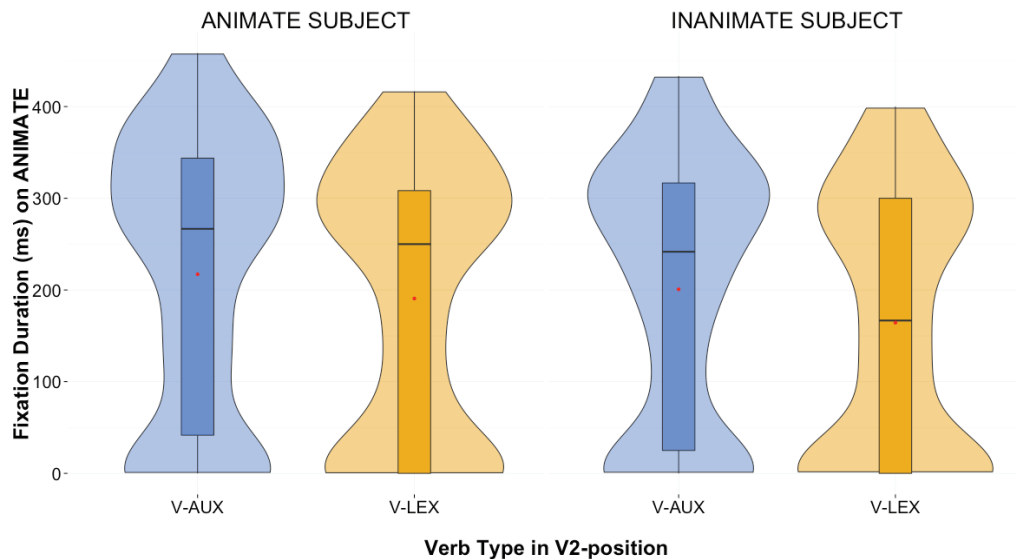
Figure 5: Violin/boxplot of fixation duration to the animate referent per condition in the early subject integration window. The red dot encodes the mean, the black vertical line encodes the median, the top and bottom edges of the box encode the interquartile range and the transparent beans provide a density plot. Note that the statistical analysis was performed on the logit-transformed odds ratio of fixations to the animate referent over those to the inanimate referent.

3.3 Subject integration window

Figure 3 suggests that the animacy preference perseveres into the subject integration window, and that it is particularly strong. Even when participants heard an inanimate subject mentioned (the straight lines) the fixation proportion did not drop below 0.5. The preference is most clearly visible for sentences with an auxiliary and an inanimate subject (straight blue line): the probability of fixating the animate starts decreasing only around 400ms after subject onset. The effect of verb type also seems to extend to the subject integration window: there seems to be a lower probability of fixating animate referents after semantically rich verbal information has become available. However, after having encountered an animate noun, looks to the animate referent seem to increase quickly in the lexical verb condition.

3.3.1 Early subject integration

Figure 5 shows the average fixation durations across conditions in the early subject integration window. Statistical analysis revealed a main effect of VERB TYPE trending towards significance (β = 0.36, SE = 0.19, p = .060), with less looks to the animate referent in sentences with a lexical verb. The main effect of ANIMACY CONFIGURATION, pointing in the direction of more looks to the animate referent in sentences with an animate subject, missed significance (β = 0.34, SE = 0.16, p = .051), The interaction was not significant (p = .32). Again, the model's intercept deviated significantly from zero (β = 1.71, SE = 0.25, p < .001), confirming the persistence of the animacy preference.

3.3.2 Late subject integration

Figure 6 shows the duration of animate fixations in the late subject integration window. Statistical analysis revealed a significant main effect of ANIMACY CONFIGURATION (β = -1.38, SE = 0.35, p = .0003), indicating there were fewer looks to the animate referent when the subject was inanimate. There was no main effect of VERB TYPE (p = .20), nor a significant interaction (p = .67). The model's intercept deviated significantly from zero (β = 1.13, SE = 0.19, p < .001), again confirming the persistence of the animacy preference.
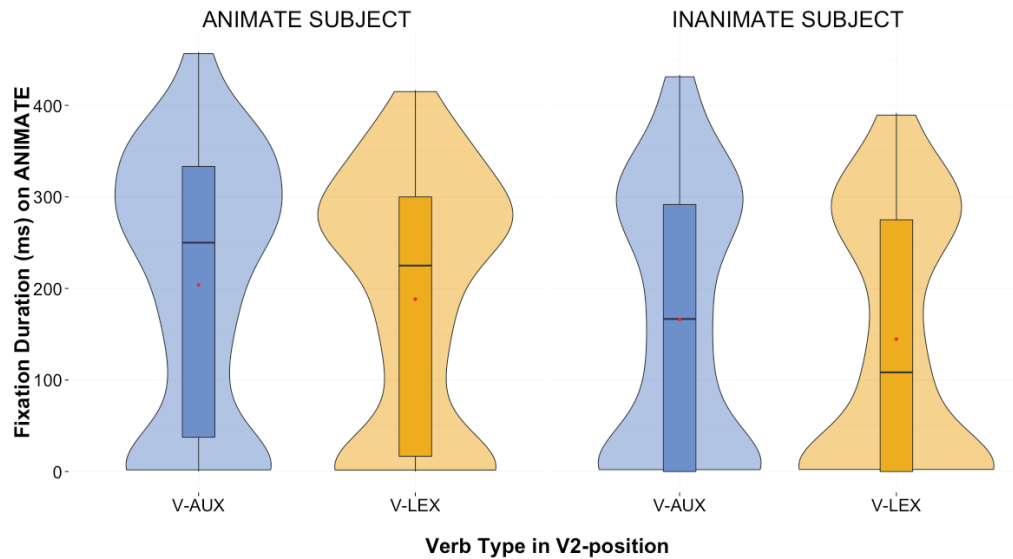
Figure 6: Violin/boxplot of fixation duration to the animate referent per condition in the late subject integration window. The red dot encodes the mean, the black vertical line encodes the median, the top and bottom edges of the box encode the interquartile range and the transparent beans provide a density plot. Note that the statistical analysis was performed on the logit-transformed odds ratio of fixations to the animate referent over those to the inanimate referent.

3.4 Object integration window

Figure 7 shows the fixation proportions to the animate over the inanimate referent in the four conditions, time-locked to object onset. The gazes during the object window contrast sharply with the patterns observed during processing of the subject NP. Figure 7 suggests that the gaze patterns closely mirror the audio input: participants focus on the animate referent in the conditions with an animate object (straight lines) and on the inanimate referent in conditions with an inanimate object (dotted lines). Verbal information seems to have an effect in the early window: for animate objects, fixations to the animate referent seem lower if the sentence contains an auxiliary, but higher when the object is inanimate.

3.4.1 Early object integration

Figure 8 shows the fixation durations in the early object integration window. Statistical analysis revealed no significant main effects of VERB TYPE (p = .65) or ANIMACY CONFIGURATION (p = .41). The interaction, reflecting that the probability of fixating the inanimate referent is higher after hearing an inanimate object compared to an animate object, but only in the lexical verb condition, was not significant (p = .070). Also, the model's intercept did not significantly deviate from zero (β = 0.14, SE = 0.21, p = .52), indicating there was no longer a bias towards the animate referent.

3.4.2 Late object integration

Figure 9 shows the fixation durations in the late object integration window. Statistical analysis revealed a significant main effect of ANIMACY CONFIGURATION (β = 3.06, SE = 0.36, p < .001), indicating there were more looks to the animate referent when the object was animate. There was no main effect of VERB TYPE (p = .21) nor a significant interaction between the two factors (p = .21). The model's intercept deviated significantly from zero (β = 0.40, SE = 0.18, p = .038), indicating more looks to the animate referent.
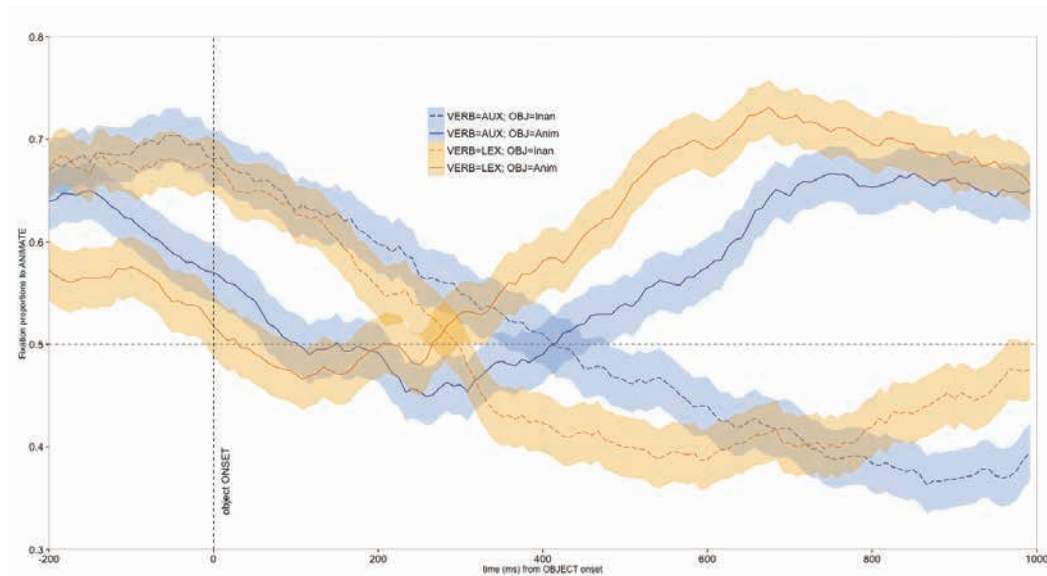
Figure 7: Gaze plot of animate fixation proportions in each condition time-locked to object onset. The lines indicate mean fixation proportions and the transparent bands indicate 1 SE above/below the means. Blue lines indicate conditions with an auxiliary finite verb, orange lines indicate conditions with a lexical finite verb, dashed lines indicate conditions with an inanimate object, and straight lines indicate conditions with an animate object.
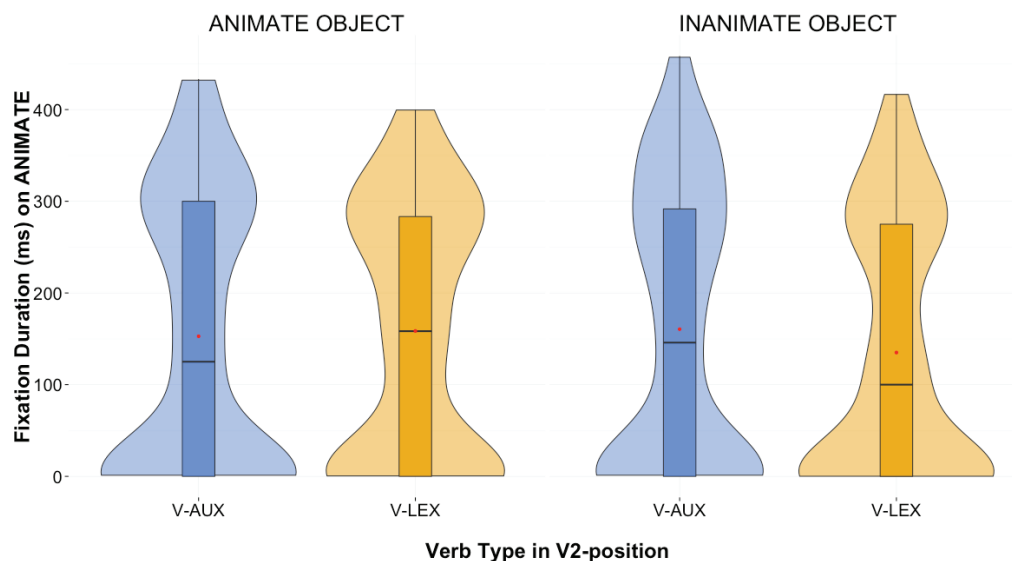

Figure 8: Violin/boxplot of fixation duration to the animate referent per condition in the early object integration window. The red dot encodes the mean, the black vertical line encodes the median, the top and bottom edges of the box encode the interquartile range and the transparent beans provide a density plot. Note that the statistical analysis was performed on the logit-transformed odds ratio of fixations to the animate referent over those to the inanimate referent.

4 Discussion

The animacy hierarchy dictates the grammaticality and interpretation of transitive sentences in many languages in the world. This hierarchy has also been argued to influence language production and comprehension in languages where it has no grammatical status. The present study set out to determine whether the animacy preference is visible in the latter type of languages during incremental sentence comprehension; moreover, it investigated whether such a preference would be affected by the lexical semantic information provided by a verb preceding the arguments in the sentence. Our results indicate that both questions can be answered affirmatively.
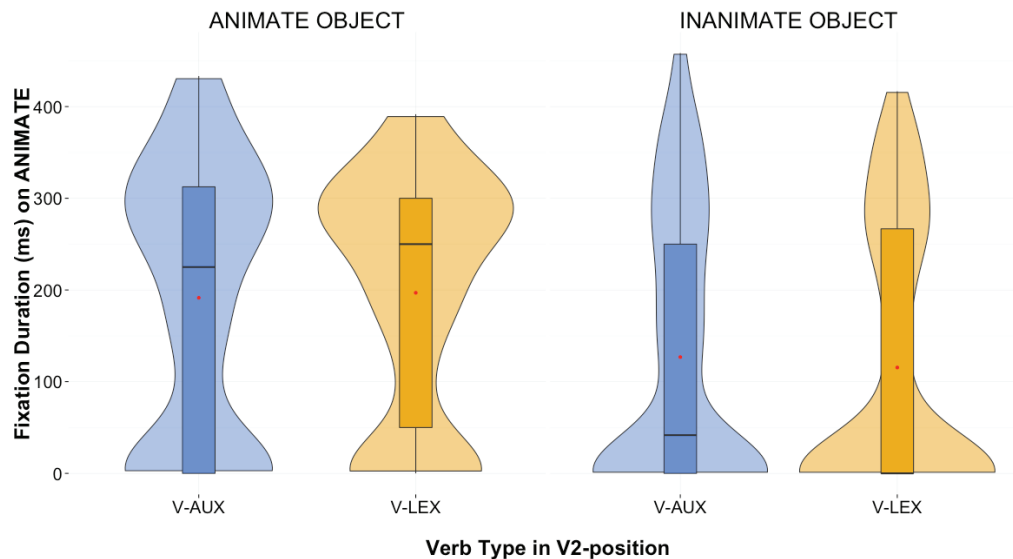
Figure 9: Violin/boxplot of fixation duration to the animate referent per condition in the late object integration window. The red dot encodes the mean, the black vertical line encodes the median, the top and bottom edges of the box encode the interquartile range and the transparent beans provide a density plot. Note that the statistical analysis was performed on the logit-transformed odds ratio of fixations to the animate referent over those to the inanimate referent.

The visual world paradigm we used rests upon a linking hypothesis, which assumes that as a consequence of utterance planning or comprehension, visual attention moves to a particular object in the visual display, resulting in a rapidly following saccadic eye movement to bring that attended area into foveal vision (Salverda & Tanenhaus 2017). This kind of looking behavior was clearly observed during object integration: people fixated object referents quickly after encountering the direct object in the speech stream. During subject integration, however, we observed a persistent preference to attend to animate over inanimate referents, even when the audio signal mentioned an inanimate argument. In fact, the preference to attend to animate over inanimate entities was already present in the pre-verbal time-window, i.e., before *any* relevant linguistic information became available. We would not want to argue that this early and persistent animacy bias is necessarily related to the incoming linguistic signal, i.e., that listeners were anticipating an animate subject (although with the current results we cannot exclude this possibility). The increased attention to animate over inanimate referents more likely reflects an *inherent* animacy bias, i.e., animate entities are more prominent, or salient, than inanimate entities (e.g. Bock & Warren 1985), regardless of their linguistic status or frequency considerations. But whether or not the animacy bias is related to predictive sentence processing, we can conclude that it is a very strong processing cue. Our results show that it was strong enough to outrank the influence of verbal information during subject integration: even when the verbal information pushed the interpretation of the subject towards the inanimate referent, the proportion of looks to the inanimate referent did not increase to more than 0.5.

We did find that verbal semantics modulated the animacy preference: we observed that in the lexical verb conditions, the proportion of animate fixations decreased quickly after the lexical information of the verb became available (in the pre-argument window). We thus corroborate earlier findings showing that lexical information in V2 position influences incremental argument interpretation (Dahan & Tanenhaus 2004, Bader & Bayer 2006). The exploratory analysis of the object integration window may provide a preliminary additional pointer to the effect of verbal information: in sentences with an animate object, the proportion of animate fixations in the early object window seemed lower when the sentence had an auxiliary verb than when it had a lexical verb. This effect was in this direction, although not significant. Speculatively, it may suggest that it is more difficult to integrate an animate object in the absence of supporting verbal information, which is a corollary of the animacy preference (dictating to assign the animate argument the subject function). This could extend the findings of Paczynski & Kuperberg (2011), who observed an increase in processing cost (i.e., more enhanced N400 amplitude) for animate vs. inanimate objects in sentences

where the lexical verb preceded the object: our findings suggest that the preceding verbal information could reduce these costs. Clearly, more research is needed to confirm this.

Assuming that listeners immediately use the thematic restrictions of the verb to anticipate an upcoming subject (cf. Dahan & Tanenhaus 2004), we hypothesized a difference between verb types within the lexical verb condition: we expected inanimate verbs to elicit more predictive looks to the inanimate referent than animate verbs. This prediction was not borne out: animate and inanimate verbs, although requiring their subjects to have opposite animacy values, elicited similar looking behavior. This suggests that listeners did not make full use of the lexical information provided by the verb to predict upcoming arguments.

A possible explanation for the absence of this interaction effect might be a more shallow parsing mechanism underlying the observed fixation patterns. It may, for instance, involve mere semantic association between the verbs and the two referents. It could well be that hearing a verb like *extinguish* will elicit more looks to a fire than to a baker, the former being more strongly associated with the meaning of the verb. Indeed, Kukona et al. (2011), also using a Visual World paradigm, found evidence for this kind of thematic priming, in addition to prediction effects based on structural characteristics. Participants were found to fixate an associated agent like *policeman* when hearing the verb *arrest*, even though a different agent had already been mentioned (*Toby arrested the crook*). If thematic priming also played a role in our experiment, e.g., if animate verbs happen to be associated more strongly with their inanimate noun, this could overshadow potential differences between verb types on predictive argument processing.

To investigate possible effects of such thematic priming, we administered two post-hoc tests to determine the association strength between the verbs and the corresponding nouns for each experimental item. Twenty-three participants (18 female, mean age 22.5 years), who did not take part in the main experiment, filled out an on-line forced choice task. They were presented with a verb and a corresponding noun pair, and had to indicate which of the nouns was most strongly associated with the verb. Each participant saw all noun pairs; verbs were counterbalanced across two lists, such that each participant saw each noun pair in combination with either the animate or the inanimate verb. Twenty-five other participants (18 female, mean age 21.8 years) filled out a rating task in which they were presented with a verb and a single noun, and had to indicate their association on a five-point Likert scale. Each participant was presented with all verbs; nouns were counterbalanced across two lists, such that each participant saw each verb in combination with either the animate or the inanimate noun. Statistical analyses of the results revealed that verbs were indeed more strongly associated with inanimate nouns than with animate nouns in both tasks. A logistic mixed-effects analysis of the results of the forced choice task showed that on average there was a significantly higher probability of associating the verb with the inanimate noun than with the animate noun (β = 1.74, SE = 0.42, p < .001), although 7 of the 32 verbs showed the opposite pattern; the odds did not significantly differ between verb types (p = 0.97). For the rating task, results from a linear mixed-effects analysis showed that on average the mean association rating between verbs and inanimate nouns was significantly higher than between verbs and animate nouns (β = 0.42, SE = 0.13, p < .001), although 8 of the 32 verbs showed the opposite pattern; again, we found no significant difference between verb types (p = 0.54), nor an interaction effect (p = 0.90).

To assess whether thematic priming was a predictor for predictive argument fixations, we reclassified the verbs into two association categories (animate vs. inanimate association) using the results of the rating task. Eight verbs were classified as showing animate association and the other 24 as inanimate association. We re-analyzed the fixations in the pre-argument window with a mixed-effects logistic regression model substituting this new factor for Animacy Configuration as a fixed effect. Given that thematic priming can only occur when the lexical verb precedes the arguments, we only analyzed trials with the lexical verb in second position. Results from this model still showed an overall preference to look at the animate referents; we found no significant differences in predictive looks between verbs that are more strongly associated with animate nouns than with inanimate nouns (p = 0.52). It should of course be kept in mind that our study was not designed to address this question, and the verb association categories are not equally divided (there were more inanimate-associated than animate-associated verbs), but the results seem to suggest that thematic priming is not the driving force behind the pre-argument fixation patterns.

An alternative explanation of the decrease in looks to the animate referent after encountering the verb could be that verb-based predictions do not concern the first upcoming argument, but *both* upcoming arguments. Upon hearing the lexical verb, the listener can start building up a representation of the corresponding transitive event (see also Sauppe 2016). As both the animate and the inanimate referent will figure in the event (be it as the actor or the undergoer), this will elicit an increase in visual attention to both referents, i.e., the proportion of animate and inanimate fixations will be more equally divided. The auxiliary *heeft* then may be semantically too empty for building up any specific event representation. Hence, a listener's visual attention will likely be drawn to the inherently most salient referent in the display.

We have shown that specific linguistic information can facilitate the integration of a globally unexpected animate object, but the lexical information from the verb alone was not sufficient to counter the inherent animacy bias. The question is whether the animacy preference can be fully countered on the basis of sufficient linguistic cues. Can we build up a linguistic context that is so strong that its effects outrank the animacy preference? Results from Szewczyk & Schriefers (2011, 2013) suggest that this is indeed possible: they show that if the predictability of an animate object on the basis of the preceding linguistic context is very high (cloze probability 99%), an inanimate object elicits enhanced processing difficulty (and the other way around). In their experiment, expectations about upcoming objects were based on much more linguistic input than just a lexical verb: participants read short stories, in which the direct object occurred as the final word of the story-final sentence. We thus seem to need ample linguistic evidence going against the global animacy bias to change our predictions about upcoming arguments (see also Muralikrishnan et al. 2015).

5 Conclusions

Our results add to the existing body of research showing the role of animacy during language comprehension. It suggests that the animacy preference in (1) is active also in languages in which it has not fully grammaticalized. At the same time, our results add to the (as of yet) limited knowledge of the effect on language processing of lexical verbal information preceding arguments.

It is well established by now that the animacy preference plays a role in language processing and in the grammar of languages around the world. It is still an open question how to model this relation between grammar and processing. The 'soft constraints mirror hard constraints' approach of Bresnan et al. (2001) provides a principled account of this relation. The constraint-based nature of this grammatical model can also easily deal with the fact that the animacy preference interacts with other principles, such as information from the verb. This model is, however, by no means the only possible approach. Alternatively, one could maintain, as envisaged by Minkoff (2000), that the animacy preference is only part of the human parser and that grammatical reflexes of this preference, as for instance found in Mam-Maya, are the result of the interaction of this parsing preference with grammatical properties of the language; in this case the fact that the language has pro-drop and is verb-initial. Which approach is more parsimonious remains to be seen. Distinguishing between these two alternative approaches is not trivial and requires a great deal of work on languages with diverse linguistic characteristics.

Abbreviations: PST – past tense; PTCP – participle; V2 – verb second

References

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21 (3). 435-448.

Altmann, Gerry. T.M., Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition* 73. 247–264.

Audacity Team. http://audacityteam.org.

Bader, Markus, Josef Bayer. 2006. *Case and linking in language comprehension: Evidence from German*. Berlin: Springer.

Bock, J. Kathryn, Richard K. Warren. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21. 47-67.

Barr, Dale J., Roger Levy, Christoph Scheepers, Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278.

van Bergen, Geertje. 2011. *Who's first and what's next: Animacy and word order variation in Dutch language production*. PhD thesis. Nijmegen: Radboud University Nijmegen.

van Berkum, Jos J.A., Colin M. Brown, Pienie Zwitserlood, Valesca Kooijman, Peter Hagoort. 2005. Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31. 443-467.

Boersma, Paul, David Weenink. 2016. Praat: doing phonetics by computer [Computer program]. Version 6.0.21, retrieved 25 September 2016 from http://www.praat.org/.

Bosker, Hans Rutger, Hugo Quené, Ted Sanders, Nivja de Jong. 2014. Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language* 75.104-116.

Bornkessel-Schlesewsky, Ina, Andrej Malchukov, Marc Richards. 2015. *Scales and hierarchies: A cross-disciplinary perspective*. Berlin: de Gruyter.

Bornkessel-Schlesewsky, Ina, Matthias Schlesewsky. 2009. The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass* 3. 19-58.

Bornkessel-Schlesewsky, Ina, Matthias Schlesewsky. 2013. Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language* 125. 60-76.

Branigan, Holly P., Martin J. Pickering, Mikihiro Tanaka. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118 (2). 172-189.

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

Cohn, Neil, Martin Paczynski. 2013. Prediction, events, and the advantage of agents: The processing of semantic roles in visual narrative. *Cognitive Psychology* 67. 73-97.

Comrie, Bernard. 1989. *Language Universals and Linguistic Theory*. Chicago: University of Chicago Press.

Czypionka, Anna. 2014. *The interplay of object animacy and verb class in representation building*. PhD thesis. Berlin: Humboldt-Universität zu Berlin, Germany.

Dahan, Delphine, Michael K. Tanenhaus. 2004. Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30. 498-513.

Dahl, Östen. 2000. Animacy and the notion of semantic gender. In Unterbeck, Barbara, Matti Rissanen, Terttu Nevalainen, Mirja Saari (eds.), *Gender in grammar and cognition I: Approaches to gender*, 99-115. Berlin: de Gruyter.

Dahl, Östen, Kari Fraurud. 1996. Animacy in Grammar and Discourse. In Fretheim, Thorstein, Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 47-64. Amsterdam/Philadelphia: John Benjamins Publishing Company.

DeLong, Katherine A., Thomas P. Urbach, Marta Kutas. 2005. Probabilistic word preactivation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* 8. 1117-1121.

Fauconier, Stefanie. 2011. Differential agent marking and animacy. *Lingua* 121 (3). 533-547.

Federmeier, Kara D. 2007. Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology* 44. 491-505.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Kamide, Yuki. 2008. Anticipatory processes in sentence processing. *Language and Linguistics Compass* 2. 647-670.

Kemmerer, David. 2012. The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's area. *Language and Linguistics Compass* 6. 50-66.

Kempen, Gerard, Karin Harbusch. 2004. A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In Pechmann, Thomas, Christopher Habel (eds.), *Multidisciplinary approaches to language production*, 173-181. Berlin: de Gruyter.

Kukona, Anuenue, Shin-Yi Fang, Karen A. Aicher, Helen Chen, James S. Magnuson. 2011. The time course of anticipatory constraint integration. *Cognition* 119 (1). 23-42.

Kuperberg, Gina R., T. Florian Jaeger. 2016. What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* 31. 32-59.

Lamers, Monique J.A. 2012. Argument linearization in Dutch: A multi-factorial approach. In Lamers, Monique J.A., Peter de Swart (eds.), *Case, Word Order, and Prominence: Psycholinguistic and Theoretical Approaches to Argument Structure,* 121-144. Dordrecht: Springer.

Lamers, Monique J.A., Helen de Hoop. 2014. Animate object fronting in Dutch: A production study. In MacWinney, Brian, Andrej Malchukov, Edith Moravcsik (eds.), *Competing Motivations in Grammar Usage.* 42-53, Oxford: Oxford University Press.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Narr.

Levinson, Stephen C. 2012. The original sin of cognitive science. *Topics in Cognitive Science* 4. 396-403.

Lockwood, Hunter T., Monica Macaulay. 2012. Prominence hierarchies. *Language and Linguistics Compass* 6. 431-446.

MacWhinney, Brian, Elizabeth Bates. 1989. *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Mak, Willem M., Wietske Vonk, Herbert Schriefers. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47. 50-68.

Mak, Willem M., Wietske Vonk, Herbert Schriefers. 2006. Animacy in processing relative clauses: the hikers that rocks crush. *Journal of Memory and Language* 54. 466-490.

Malchukov, Andrej 2008. Animacy and asymmetries in differential case marking. *Lingua* 118. 203-221.

Marslen-Wilson, William D. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25. 71-102.

Minkoff, Seth. 2000. Animacy hierarchies and sentence processing. In Carnie, Andrew, Eithne Guilfoyle (eds.), *The syntax of verb initial languages*, 201-209. Oxford: Oxford University Press.

Muralikrishnan, R., Matthias Schlesewsky, Ina Bornkessel-Schlesewsky. 2015. Animacy-based predictions in language comprehension are robust: contextual cues modulate but do not nullify them. *Brain Research* 1608. 108-137.

Norcliffe, Elisabeth, Alice C. Harris, T. Florian Jaeger. 2015. Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language, Cognition and Neuroscience* 30. 1009-1032.

Ortmann, Albert. 1998. The Role of [+/-animate] in Inflection. In Fabri, Ray, Albert Ortmann, Teresa Parodi (eds.), *Models of inflection*, 60-84.Tübingen: Niemeyer.

Øvrelid, Lilja. 2004. Disambiguation of syntactic functions in Norwegian: Modeling variation in word order interpretations conditioned by animacy and definiteness. In Karlsson, Fred (ed.), *Proceedings of the 20th Scandinavian Conference of Linguistics*. University of Helsinki: Department of General Linguistics.

Paczynski Martin, Gina R. Kuperberg. 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.

Pickering, Martin J., Simon Garrod. 2013. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences* 36. 329-347.

Primus, Beatrice. 2012. Animacy, generalized semantic roles, and differential object marking. In Lamers, Monique J.A., Peter de Swart (eds.), *Case, word order, and prominence: Interacting Cues in Language Production and Comprehension*, 65-90. Dordrecht: Springer.

Psychology Software Tools (2009). E-prime2 with the extensions for Tobii. https://www.pstnet.com/downloads/eet/ EETVersion6.pdf.

Salverda, Anne Pier, David Kleinschmidt, Michael K. Tanenhaus. 2014. Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of Memory and Language* 71 (1). 145-163.

Salverda, Anne Pier, Michael K. Tanenhaus. 2017. The visual world paradigm. In de Groot, Anette M.B., Peter Hagoort (eds.), *Research Methods in Psycholinguistics*, 89-110. Hoboken, NJ: Wiley.

Sauppe, Sebastian. 2016. Verbal semantics drives early anticipatory eye movements during the comprehension of verb-initial sentences, *Frontiers in Psychology* 7.

Scheepers, Christoph, Barbara Hemforth, Lars Konieczny. 2000. Linking syntactic functions with thematic roles: Psych-verbs and the resolution of subject-object ambiguity. In Hemforth, Barbara, Lars Konieczny (eds.), *German sentence processing*, 95-135. Dordrecht: Kluwer.

Siemund, Peter. 2008. *Pronominal gender in English: A study of English varieties from a cross-linguistic perspective*. London: Routledge.

Siewierska, Anna. 1988. *Word order rules*. London: Croom Helm.

Snider, Neal, Annie Zaenen. 2006. Animacy and syntactic structure: fronted NPs in English. in Butt, Miriam, Mary Dalrymple, Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*, 323-338. Stanford: CSLI Publications.

de Swart, Peter. 2007. *Cross-linguistic variation in object marking*. Utrecht: LOT Publications.

Szewczyk Jakub M., Herbert Schriefers. 2011. Is animacy special? ERP correlates of semantic violations and animacy violations in sentence processing. *Brain Research* 1368. 208-22.

Szewczyk Jakub M., Herbert Schriefers. 2013. Prediction in language comprehension beyond specific words: An ERP study on sentence comprehension in Polish. *Journal of Memory and Language* 68. 297-314.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, Julie C. Sedivy. 1995. Integration of visual and linguistic information during spoken language comprehension. *Science* 268. 1632-1634.

Tomlin, Russell S. 1986. *Basic word order: Functional principles*. London: Croom Helm.

Trueswell, John C., Michael K. Tanenhaus, Susan M. Garnsey. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33. 285-318.

Van Petten, Cyma, Barbara J. Luka. 2012. Prediction during language comprehension: Benefits, costs and ERP components. *Journal of Psychophysiology* 83. 176-190.

Van Valin, Robert D., Randy J. LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge University Press, Cambridge.

Yamamoto, Matsumi. 1999. *Animacy and reference: A cognitive approach to corpus linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Young, Robert W., William Morgan. 1987. *The Navajo language: A grammar and colloquial dictionary*. Albuquerque: University of New Mexico Press.

Wu, Fuyun, Elsi Kaiser, Elaine Andersen. 2012. Animacy effects in Chinese relative clause processing. *Language And Cognitive Processes* 27. 1489-1524.

APPENDIX A

List of noun-verb combinations used in the stimuli with their English translations. The label '(in)animate verb' indicates that given the combination with the two nouns, the (in)animate noun is the one taking up subject position. Although many verbs could in principle take both animate and inanimate nouns as either their subject or object, the specific combination with the selected noun pairs left only one plausible configuration per verb. The label should thus not be taken to imply that the given verb cannot occur with an (in)animate subject at all.

	Inanimate noun	Animate noun	Inanimate verb	Animate verb
1	*camera*	*dief*	*registreren*	*stelen*
	'camera'	'thief'	'to record'	'to steal'
2	*roman*	*koning*	*ontroeren*	*lezen*
	'novel'	'king'	'to move, touch'	'to read'
3	*tekening*	*bisschop*	*grieven*	*verscheuren*
	'drawing'	'bishop'	'to hurt, offend'	'to tear (up)'
4	*medaille*	*zwemmer*	*motiveren*	*ontvangen*
	'medal'	'swimmer'	'to motivate'	'to receive'
5	*waterval*	*fotograaf*	*inspireren*	*vereeuwigen*
	'waterfall'	'photographer'	'to inspire'	'to immortalize (e.g. on film)'
6	*sneeuw*	*piloot*	*hinderen*	*trotseren*
	'snow'	'pilot'	'to impede'	'to defy'
7	*explosie*	*politieman*	*doden*	*voorkomen*
	'explosion'	'policeman'	'to kill'	'to prevent'
8	*brand*	*bakker*	*duperen*	*blussen*
	'fire'	'baker'	'to dupe/victimize'	'to extinguish (a fire)'
9	*sluier*	*tandarts*	*benadelen*	*weigeren*
	'veil'	'dentist'	'to put at a disadvantage'	'to refuse'
10	*piano*	*verhuizer*	*verpletteren*	*verplaatsen*
	'piano'	'mover'	'to crush'	'to move'
11	*golf*	*surfer*	*overspoelen*	*gebruiken*
	'wave'	'surfer'	'to wash over'	'to use'
12	*bliksem*	*visser*	*treffen*	*zien*
	'lightning'	'fisherman'	'to hit'	'to see'
13	*kerkklok*	*zwerver*	*wekken*	*horen*
	'church bell'	'tramp'	'to wake'	'to hear'
14	*kogel*	*militair*	*verwonden*	*ontwijken*
	'bullet'	'soldier'	'to wound'	'to dodge'
15	*ketting*	*bruid*	*verrassen*	*bewonderen*
	'necklace'	'bride'	'to surprise'	'to admire'
16	*sjaal*	*kleuter*	*beschermen*	*dragen*
	'scarf'	'pre-schooler'	'to protect'	'to wear'