Verb type, animacy and definiteness in grammatical function disambiguation*

Monique J.A. Lamers
Radboud University/CLS

1. Introduction

To build up a meaningful construction, the speaker and hearer have several sources of information (including grammatical rules, situational knowledge, and general world knowledge) at their disposal. This paper addresses the role of different sorts of information in the combinatory process of the hearer. More specifically, with a comprehension study, I will show that animacy and definiteness diverge in the interplay with three different types of verbs.

Firstly, I will briefly discuss three sources of information available in an utterance. The three sources of information are (1) verb type and (2) animacy, known to be essential for argument comprehension, and (3) definiteness, which might also play an important role. It will be discussed how they relate to two well-known word order preferences in argument structure, namely the so-called subject-first and animate-first preference. Section 3 presents a rating study that was performed to investigate what the exact interplay is between these three sources of information. To explain the results, in Section 4 I will propose an analysis in which the hearer takes into account the possibilities of the speaker to use a form in which the subject-first and animate-first preference are fulfilled. Finally, I will draw some conclusions (Section 5).

2. Argument structure: Verbs, arguments and word order

2.1 Verbs, arguments and word order

Verbs come with their own argument structure. They establish the syntactic and semantic relationship between the arguments, representing the participant(s) of the event or state expressed by the verb. In the argument structure of a verb semantic roles (e.g. agent, patient, experiencer) are matched with syntactic functions.
Arguments have to meet certain criteria, the selectional restrictions of the verb, to fit the conceptual structure of the verb. For instance, the verb \textit{admire} needs two arguments, the \textit{admirer} (someone doing the admiring), and the \textit{admiree} (the entity being admired). Given the meaning of \textit{admire}, the admirer must have animate or human characteristics, whereas no such constraints apply to the admiree:

(1) The athlete admired the picture.

Because of the strict word order in English, the subject is almost always the first semantic argument of an active sentence. If the verb \textit{admire} would be replaced by \textit{please}, as in (2), we cannot make sense of this sentence.

(2) The athlete pleased the picture.

Although \textit{please} and \textit{admire} both take two arguments (subject and object), the semantic argument structure is different: \textit{please} does not constrain the subject argument, but rather the object, which must be able to experience the psychological state expressed by this verb.

Hence, especially for languages with a free word order, the animacy of the arguments provides important information to derive the correct structure and meaning of an utterance. For example, in Dutch, a language with no case marking on full NPs and a relatively free word order, main clauses have to be disambiguated in a subject-before-object (SO) or object-before-subject (OS) structure. Because of a strong preference for SO structures over OS structures, sentences in which no disambiguating information is provided, as in (3a), will be interpreted as SO. This structural preference, dubbed as the subject-first preference, is found in comprehension as well as in production resulting in differences in lower processing costs and higher occurrences for the preferred structure (e.g., Frazier & Flores d’Arcais 1989; Lamers 2001, 2005; Bock & Warren 1985). From a theoretical perspective it is argued that subject-first sentences are less complex, following parsing mechanisms such as the active filler strategy resulting in a parsimonious structure (cf. Frazier & Flores d’Arcais 1989; De Vincenzi 1991). In contrast to (3a), however, in (3b) and (3c) it is the animacy in combination with the selectional restrictions of the verb that resolves the ambiguity.

(3) a. De fotograaf verwonderde/beviel de atleet. \hspace{1cm} <SO/OS>
    the photographer\textsubscript{SBJ/OBJ} amazed/pleased the athlete\textsubscript{OBJ/SBJ}
    ‘The photographer amazed/pleased the athlete.’

b. De foto verwonderde/beviel de atleet. \hspace{1cm} <SO>
    the picture amazed/pleased the athlete
    ‘The picture amazed/pleased the athlete.’
c. De atleet verwonderde/beviel de foto.
the athlete amazed/pleased the picture
‘The picture amazed/pleased the athlete.’

Since animate NPs are good agents and experiencers, it is not surprising that most subjects are animate (cf. Primus 1999; Dowty 1991). Moreover, there are numerous verbs assigning the role of agent or experiencer to the subject. Corpus research corroborates this observation: in a corpus study of spoken Swedish about two thirds of all subjects are animate (Dahl 2000; for similar observations in Norwegian, see Øvrelid, 2004). Kempen and Harbus (2004) also found a high incidence of animate subjects in German, a language with a free word order and case morphology on full NPs. They report a direct influence of animacy on word order irrespective of grammatical function with a clear preference of structures with an animate NP as the first NP of the sentences, i.e., a pure animate-first effect. There is also ample evidence in psycholinguistic studies for the preference for structures in which an animate NP precedes an inanimate one (for production see Bock and Warren 1985, and Prat-Sala and Branigan 2000; for comprehension see Lamers 2001, 2005 and Weckerly and Kutas 1999). For prototypical transitive verbs, as well as for subject-experiencer verbs, such as admire, this conceptual animate first preference goes hand in hand with the structural subject-first preference. This is however not the case for verbs as verwonderen ‘amaze’ and bevallen ‘please’ (see examples (2) and (3)). In Dutch, these two verbs belong to two different classes of psych verbs, causative psych verbs (a.k.a. object-experiencer verbs), and unaccusative psych verbs. These two types of psych verbs have similar syntactic frames with a subject and object argument, and similar selectional restrictions, constraining the object argument to animate NPs. Hence, if these verbs are combined with an animate and inanimate NP, the animate-first preference is not fulfilled if the structure is subject-initial, whereas if the animate first preference is fulfilled in object-initial sentences, the subject-first preference is not fulfilled ((3b) and (3c), respectively). However, for causative psych verbs a (pseudo)passive construction is allowed fulfilling both preferences, as is illustrated in (4a). Since unaccusative psych verbs cannot passivize, there is no form available that can fulfil the two preferences simultaneously (4b).

(4) a. De atleet werd verwonderd door de foto.
the athlete was amazed by the picture
‘The athlete was amazed by the picture.’

b. *De atleet werd bevallen door de foto.
the athlete was pleased by the picture
‘The athlete was pleased by the picture.’
The verb types (agentive/subject-experiencer verbs, causative psych verbs and unaccusative psych verbs) described in this section, are used in a rating study that will be presented in Section 3. Given the differences between the verb types in relation to the word order preferences, differences in comprehension can be expected between structures in which different word orders and different types of verbs are used. Verb type and animacy are, however, not the only factors that might influence the comprehension process. As will be discussed in the next section, definiteness might also play an important role.

2.2 Animacy and definiteness

Whereas animacy refers to an animate or inanimate entity, definiteness refers to the identifiability of the referent, being definite (and specific), or indefinite (and specific, non-specific or general). In Dutch, definiteness is marked by different articles with *de* ‘the’ and *het* ‘the’ for definites, and *‘n* or *een* ‘a’ for indefinites. In contrast to animacy, which refers directly to properties of the (discourse) referents (i.e., the individuals the noun phrases refer to), definiteness does not reflect an inherent property of the discourse referent; it merely reflects its role or status in the discourse (it is supposed to be unique or familiar in the discourse (cf. Lambrecht 1994)). In addition, the interpretation of definiteness might depend on the position in the sentence, as is illustrated in (5). In (5a) the indefinite second argument *een atleet* ‘an athlete’ gets a non-specific-reading, whereas in (5b) at the sentence initial position, it gets a specific reading.

(5)  

a. De foto beviel een atleet.  
   *the picture pleased the athlete*  
   ‘The picture pleased an athlete.’

b. Een atleet beviel de foto.  
   *an athlete pleased the picture*  
   ‘The picture pleased an athlete.’

In the literature, animacy and definiteness are sometimes treated as being similar with respect to grammatical function. For instance, corpus studies show that subjects are prototypically definite and animate whereas objects are likely to be indefinite and inanimate (Dahl and Fraurud 1996). In languages with differential object marking only definite and animate direct objects, both referring to entities that are high in prominence, are overtly case marked whereas indefinite, inanimate ones, which are usually placed low on scales of prominence, stay unmarked (Bossong 1985; Aissen 2003).

Thus, although it can be argued that animacy and definiteness are different in referential characteristics, they also share similarities in relation to prominence.
The question arises to what extent the definiteness of the arguments influences the comprehension process. In contrast to the influence of animacy, so far the role of definiteness is hardly investigated in psycholinguistics, let alone the interplay between animacy, definiteness and verb type. The comprehension study, described in the next section, addresses this paucity.

3. Rating study: Ease of comprehensibility

3.1 Method

3.1.1 Participants
Eighty native speakers of Dutch (18–28 years old; 56 female participants) participated in the rating study. They were paid for their participation (8 euros).

3.1.2 Materials
In the comprehension study in Dutch, embedded clauses with a local structural ambiguity and three different types of verbs were used. The verb types were agentive/experiencer-theme verbs that can passivize and select an animate object (e.g. begrijpen ‘understand’ or bewonderen ‘admire’), causative psych verbs that also passivize but select an animate object (e.g. verwonderen ‘amaze’), and unaccusative psych verbs that cannot passivize and select an animate object (e.g. bevallen ‘please’). The two NPs were either two animate NPs or an animate and an inanimate NP. They could be definite (with the definite article de ‘the’) or indefinite (with the indefinite article n, which is the unambiguous clitic form of the indefinite article een ‘a’an’ and cannot be interpreted as the numeral one). Presenting them in different word orders resulted in 36 conditions. An overview of the conditions is provided in Table 1.

NPs were combined in such a way that they formed equally plausible sentences with one of the items of each verb type as was attested in a pretest. The verbs in such a set were combined with three different NP combinations of two animate NPs; one of the animate NPs of each of these combinations was combined with an inanimate NP, thus forming three animate-inanimate NP combinations. Since in Dutch there are only 8 unaccusative verbs that can be combined with two animates NPs, there were only 8 sets of 3x36 sentences. Ten additional agentive/subject-experiencer and causative psych verbs were combined with another set of NP combinations, forming 10 sets of 3x24 sentences. The total of 1584 sentences were divided over 8 lists in such a way that all NPs and verbs occurred equally often on each list as well as all conditions were equally divided over lists. The sentences on each list were presented in pseudorandomized order. To prevent an order effect the pages with approximately 20 sentences were pseudorandomly
mixed. Repetitions were kept as far apart as possible. On each lists 20 filler sentences were added that were difficult or impossible to comprehend (e.g., Dat de zanger de stalker zong… ‘That the artist the stalker sang…’, Dat de reiziger de trein treuzelde …’That the traveller the train lingered…’). These filler sentences served as a control to check whether participants paid attention in rating the sentences.

3.1.3 Procedure
Participants were asked to rate SO and OS embedded clauses on the ease of comprehensibility on a scale form 1 to 7 (1 = very difficult to comprehend, 7 = easy to comprehend). A difference in rating is taken as an indication for a difference in the comprehension process. By presenting some examples with an explanation at the beginning of a list, it was made sure that participants knew what was meant with ease of comprehensibility.

3.2 Results
All participants rated all filler sentences as extremely difficult to understand (with a rating of 1 or 2), indicating that they filled out the form seriously.
Figure 1 shows the mean ratings for experimental SO and OS embedded clauses with the three types of verbs. The results show a complex pattern of interactions between the different factors. All embedded clauses with two animate NPs were rated relatively high. Following the strong preference of SO sentences over OS it is assumed that they were interpreted as SO structures (see also Zeevat 2004, de Hoop and Lamers 2006 on word order freezing). An important finding, in line with the results of Lamers (2001), was a higher rating for SO structures in comparison to OS structures for all three types of verbs, with the highest ratings for SO structures with agentive/subject-experiencer verbs, followed by causative psych verbs and unaccusative psych verbs. Remarkably, for OS structures relatively high ratings were found for unaccusative psych verb structures. OS structures with agentive/subject-experiencer verbs were rated as most difficult to understand, closely followed by SO structures with causative psych verbs. Note that despite the use of implausible filler sentences, the rating for the OS structures with agentive/subject-experiencer verbs was very low, whereas all SO sentences were rated high on the scale of comprehensibility. This indicates that participants made use of the whole scale in their ratings of the experimental sentences making it in principle possible to track down more subtle differences in comprehensibility (as are reported below).

The data were statistically evaluated using mean ratings computed for each condition. An overall repeated measures of analysis of variance (ANOVA) was performed with Word order (SO with two animate NPs, SO with an animate and inanimate NP, SO with an animate and inanimate NP), Definiteness (definite-definite NPs, definite-indefinite NPs, indefinite-definite NPs, indefinite-indefinite NPs) and Verb type (agentive/subject-experiencer, causative psych, unaccusative psych verbs) as the within-subject factors. Because a three way interaction was found $F(12, 948) = 2.39, p < .019$), separate analyses were performed per verb type.

**Agentive/Subject-Experiencer verbs**
The ANOVA with Word order and Definiteness showed a main effect of Word order ($F(2, 78) = 1179.11, p < .001$), but not for Definiteness, nor was there a significant interaction. Pairwise comparisons showed no difference in ratings between SO structures with two animate NPs and an animate and inanimate NP; both SO structures were rated differently from OS constructions, with higher ratings for SO structures (SO with two animate vs. OS: $T(79) = 36.40, p < .01$; SO animate-inanimate vs. OS $T(79) = 34.67, p < .01$, see also Figure 1a).

**Causative psych verbs**
A similar pattern as was found for agentive/subject-experiencer verbs was found for causative psych verbs: the ANOVA with Word order and Definiteness showed
Figure 1. Overview of the mean ratings (and standard deviations) for each condition. A. Structures with agentive/subject-experiencer verbs; B. with causative psych verbs; C. with unaccusative psych verbs (SO = subject initial; OS = object initial; A = animate NP; I = inanimate NP; Def. = definite; Indef.=indefinite).
a main effect of Word order \((F(2, 78) = 864.60, p < .001)\), but not for Definiteness, nor was there a significant interaction. For this verb type, however, all pairwise comparisons showed differences in ratings with the highest ratings for structures with two animate NPs, followed by SO structures with an inanimate and an animate NP, and the lowest rating for the OS structures (SO, animate-animate vs. SO, inanimate-animate \(T(79) = 30.38, p < .01\); SO animate-animate vs OS: \(T(79) = 7.40, p < .01\); SO animate-inanimate vs. OS \(T(79) = -30.09, p < .01\), see also Figure 1b).

**Unaccusative psych verbs**

The ANOVA for unaccusative psych verbs showed significant main effects (Word order: \(F(2, 78) = 13.24, p < .001\); Definiteness: \(F(3, 77) = 11.83, p < .001\), as well as a Word order x Definiteness interaction \(F(6, 74) = 2.73, p < .023\). To follow up on this interaction separate analyses were performed for the different word orders. Remarkably, only the SO constructions showed significant effects of Definiteness (SO, animate-animate: \(F(3,77) = 9.73, p < .001\); SO, inanimate-animate: \(F(3,77) = 7.15, p < .001\)). The differences in ratings between OS constructions with different definiteness combinations were not significant. As can be seen in Figure 1c, for both SO constructions lowest ratings were found for sentences with an indefinite object argument.

In summary, the overall statistical analyses revealed an interaction of Verb type, Word order, and Definiteness. There was only a main effect of Word order, and no interaction, nor a main effect of Definiteness for structures with agentive/subject-experiencer verbs and causative psych verbs. For structures with unaccusative psych verbs an interaction of Word order and Definiteness was found, as well as main effects for these two factors.

4. **Discussion**

In the comprehensibility rating study presented in this paper, three factors (verb type, animacy, and definiteness) were manipulated. The three different types of verbs used in the experiment were agentive/subject-experiencer verbs that selected an animate subject and can passivize, causative psych verbs, that also can passivize but select an animate object, and unaccusative psych verbs that cannot passivize. Embedded clauses with an animate and an inanimate NP were disambiguated in SO or OS structures, whereas the embedded clauses with two animate NPs were in principle ambiguous. However, the high ratings for these sentences indicate that they were interpreted SO structures. A NP was either definite or indefinite.

Results of the rating study showed that verb type, word order of the animate and inanimate arguments, as well as definiteness all influence the comprehension
process. Across the board, SO structures are easier to comprehend than OS structures. But, largest difference in rating between these two structures was found with agentive/subject-experiencer verbs. Because of the selectional restrictions of agentive verbs, the SO structure is much better than the OS structure, as it fulfils both the subject-first and the animate-first preference, while the OS structure violates both when the object is inanimate.

The two preferences do not go hand in hand for the two types of psych verbs, where either the subject-first or the animate-first condition gets violated when the object is animate and the subject is inanimate. Hence, the SO structure with an inanimate and animate NP and a causative psych verb violates animate-first, resulting in a lower rating. The SO structure with two animate NPs fulfils both preferences. Yet, according to the ratings and statistical analysis, there is still a difference in ease of comprehensibility between these SO structures and those with an agentive/subject-experiencer verb. A possible explanation can be that the harmony value of the form corresponds to the rating value, such that the rating not only depends on what the optimal form is (because then we would expect the highest rating for all optimal forms) but what the actual harmony value of the form is (that is, how many constraints are violated and to what degree) (Smolensky & Legendre 2006).

The unaccusative psych verbs get by far the highest rating for the OS structures compared to the other verb types. The ratings indicate that for these verbs SO and OS structures are both easy to understand. Note that the selectional restrictions of psych verbs are the same for causative and unaccusative psych verbs. Therefore, the difference between unaccusative psych verbs and the other two types of verbs cannot be explained by selectional restrictions alone; another factor must play a role as well. I claim that this factor is the fact that passivization is not possible with unaccusative psych verbs. The small difference in rating between the subject-first SO and the animate-first OS structure is caused by the absence of a passive form that can fulfil both preferences, thus diminishing the difference in ease of comprehensibility of the two possible forms. Hence, for unaccusative psych verbs but not for causative psych verbs, the OS structure is the only form available to the speaker to satisfy animate-first. As a native language user, the hearer takes into account the possibilities of the speaker, resulting in an increase in rating of the OS structure by the hearer.

Clearly, for causative verbs, as for agentive/subject-experiencer verbs, structural word order and animacy word order provide strong cues in the comprehension process, possibly leading to optimal forms. Finding no effect of definiteness is taken as evidence that definiteness is not a strong enough cue to influence the ease of comprehensibility. For unaccusative verbs, with no possibility to fulfil both preferences simultaneously, these cues seem to be less strong which apparently
enlarges the influence of definiteness as a cue to assign the arguments to the appropriate syntactic function.

5. Conclusion

The study reported in this paper shows an influence of verb type, animacy, and definiteness in comprehension. The results show a clear preference for subject-first structures, but this preference differs in strength for three different types of verbs. I argued that the differences should be attributed to the possibility to fulfil both a structural subject-first preference and a conceptual animate-first preference. Strikingly, the influence of definiteness in the ease of comprehensibility is limited to structures with unaccusative psych verbs. For these verbs there is no way both subject-first and animate-first can be fulfilled once there is an inanimate NP involved, and this seems to open up the possibility for definiteness to play a role in the comprehension of argument structure.

Notes

* This research was supported by the Netherlands Organisation for Scientific Research. I am grateful to an anonymous reviewer, my colleagues from the Optimal Communication group (Radboud University Nijmegen) and the audience at the TIN-dag 2007 for useful comments.

1. Here ‘speaker’ will be used for the language user that produces an utterance independent of the modality (e.g., spoken or written); the term ‘hearer’ will be used for the person comprehending an utterance.

2. In contrast to Dutch, English has no unaccusative psych verb class with two arguments.

3. One of the few studies in which definiteness is addressed is the study of Kaan (1998) in which she shows that strength of word order preferences of wh-phanes varies with the NP-type of the second argument.

4. An embedded structure was chosen to be able to compare the results of this study to other studies in which similar sets of embedded clauses were used (see Lamers, 2001; Lamers, Hagoort & de Hoop, 2006).

5. For example, participants judged how likely it is that a question amazes a politician. As in the experiment a scale from 1(not likely at all)–7(very likely or plausible) was used. Care was taken that sentences with two animate NPs were disambiguated by number agreement.

6. Due to space limits, only the most relevant results and analyses are presented.

7. The Huynh-Feldt epsilon correction was used, when evaluating effects with more than one degree of freedom in the numerator to adjust for sphericity violations. The original degrees of freedom and the corrected p-values are reported.
8. Because of the difference in selectional restrictions between the agentive/subject-experiencer verbs on the one hand and the two types of psych verbs on the other hand, resulting in different animacy word orders for subject-initial and object-initial structures. By testing Syntactic structure with three levels, the conditions with two animate NPs are distinguished from the conditions with an animite and inanimate NP. Recall that the latter condition has an animate-inanimate word order in SO structures, whereas as with causative and unaccusative psych verbs the word order for SO structures is inanimate-animate. ANOVAs with *Animacy word order* as a within subject factor instead of Syntactic structure were also performed. Since these analyses did not reveal any additional insights, I decided not to report them.

9. The statistical analyses for this word order showed a significant Verb type x Definiteness interaction (F(6,74) = 4.46; p < .002). In statistical analysis in which the structures of agentive/subject-experiencer verbs and causative psych verb were compared only showed a significant main effect of Verb type (F (1,79) = 13.69, p < .001).

**References**


