

Who or what we talk about: A comparative corpus study on the animacy structure of transitive sentences and sentence subjects

Monique Lamers & Bob van Tiel

VU University Amsterdam

Radboud University Nijmegen



Goal

Investigate the effects of register, target audience, and discourse topic on the animacy distribution in Dutch and English written and spoken sentences in different types of texts.

Background and previous findings

Spoken and written language is pervasively marked by considerations of animacy. For example, in many languages subjects tend to be animate (e.g. Clark 1965).

For transitive sentences in Swedish, Dahl & Fraurud (1996) report the following pattern:

		Object	
		Animate	Inanimate
Subject	Animate	8.9	47.7
	Inanimate	2.6	40.8

With 47.7%, A/I structures are most frequent and thus called *unmarked structures* (D&F, 1996). But why is there still a prevalence of 52.3% of marked structures?

Hypothesis: The use of marked and unmarked structures depends on various corpus and discourse properties.

STUDY I

Register and target audience in written Dutch

We expect unmarked structures to occur more often (1) in texts written for children than in texts written for adults, and (2) in literary texts than in news texts.

We extracted 200 transitive main clauses from the following corpora:

	Children	Adults
Literary		
Newspaper		

A multinomial regression analysis indeed yielded main effects of Target audience and Register, as well as an interaction between the two. In each corpus, A/I occurred most often. This preference was the strongest in literary texts for children and the weakest in news texts for adults, the corpus that resembles D&F's corpus most closely. The frequencies of the different marked structures varied between corpora.

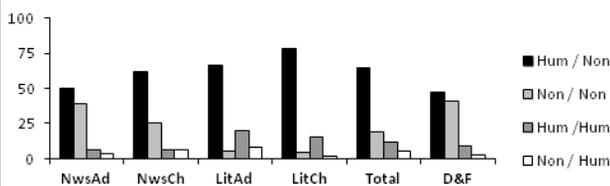


Fig. 1. Occurrences of the four structures (%) in the different corpora.

References: Clark, H.H. (1965). Some structural properties of simple active and passive sentences. *J of Verbal Learning and Verbal Behavior*, 4, 365-370; Dahl, O., & K. Fraurud (1996). Animacy and grammar in discourse. In T. Fretheim & J.K. Gundel, eds., *Reference and referent accessibility*. Amsterdam: John Benjamins, 47-64.

STUDY II

Target audience and discourse topic in English and Dutch spoken news

We expect more animate subjects in news items with an animate discourse topic, as well as more animate subjects in children's news.

We further expect an effect of language-specific characteristics. In Dutch, animacy is the strongest cue to identify the subject, whereas in English it is word order. This might result in a higher number of animate subjects in Dutch texts.

Transitive sentences were selected from 40 news items (20 with an animate and 20 with an inanimate discourse topic) taken from the following corpora:

	Children	Adults
Dutch		
English		

A binomial regression analysis revealed a main effect of Language and Topic and an interaction between Language and Target audience. There were more animate subjects than inanimate subjects in all corpora. For Dutch but not for English, inanimate subjects were less likely to occur in children's news (see Fig. 2).

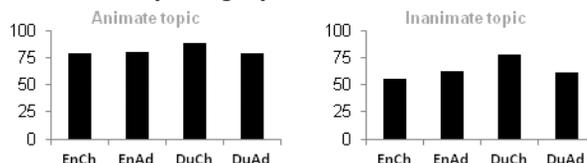


Fig. 2. Animate subjects (%) in transitive sentences in news items with an animate or inanimate discourse topic.

Taking into account all sentences, a binomial regression analysis revealed main effects of Language, Topic and Target audience, as well as a three-way interaction between the three. In news items with inanimate discourse topics, there were more inanimate subjects than animate ones. For Dutch but not for English, inanimate subjects were less likely to occur in inanimate news topics in children's news than in adults' news (see Fig. 3).

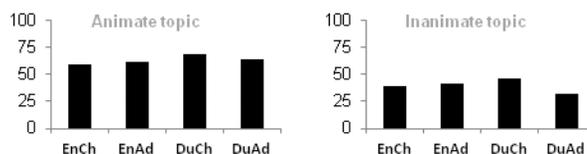


Fig. 3. Animate subjects (%) in all sentences in news items with an animate or inanimate discourse topic.

Conclusion

The distribution of marked and unmarked sentences depends on corpus characteristics that extend beyond the sentence level. These include properties like target audience and language. The animacy of the subject of a sentence also depends on what or who we talk about.

The preference for unmarked transitive sentences in all corpora reconfirms the claim of D&F (1996) that A/I structures are unmarked.

Differences between English and Dutch corpora on the one hand, and adults' news and children's news on the other hand, indicate that the use of animate subjects depends on the importance of animacy as a cue to identify the subject.

We thank Mindy Au for collecting the data of study II.
Request for reprints: monique.lamers@vu.nl; bobvantiel@gmail.com