Language Sciences 69 (2018) 30–42

Contents lists available at ScienceDirect

Language Sciences

journal homepage: www.elsevier.com/locate/langsci

A competitive mechanism selecting verb-second versus verb-final word order in causative and argumentative clauses of spoken Dutch: A corpus-linguistic study

Gerard Kempen [a,*], Karin Harbusch [b]

[a] Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[b] Faculty of Computer Science, University of Koblenz-Landau, Universitätsstraße 1, 56070 Koblenz, Germany

ARTICLE INFO

Article history:
Received 16 July 2017
Received in revised form 13 January 2018
Accepted 4 May 2018
Available online 19 May 2018

Keywords:
Causative sentences
Argumentative sentences
Verb-second
Verb-final
Dutch language
Corpus linguistics

ABSTRACT

In Dutch and German, the canonical order of subject, object(s) and finite verb is 'verb-second' (V2) in main but 'verb-final' (VF) in subordinate clauses. This occasionally leads to the production of noncanonical word orders. Familiar examples are causative and argumentative clauses introduced by a subordinating conjunction (Du. *omdat*, Ger. *weil* 'because'): the *omdat/weil*-V2 phenomenon. Such clauses may also be introduced by coordinating conjunctions (Du. *want*, Ger. *denn*), which license V2 exclusively. However, *want/denn*-VF structures are unknown. We present the results of a corpus study on the incidence of *omdat*-V2 in spoken Dutch, and compare them to published data on *weil*-V2 in spoken German. Basic findings: *omdat*-V2 is much less frequent than *weil*-V2 (ratio almost 1:8); and the frequency relations between coordinating and subordinating conjunctions are opposite (*want* >> *omdat*; *denn* << *weil*). We propose that conjunction selection and V2/VF selection proceed partly independently, and sometimes miscommunicate—e.g. yielding *omdat/weil* paired with V2. *Want/denn*-VF pairs do not occur because *want/denn* clauses are planned as autonomous sentences, which take V2 by default. We sketch a simple feedforward neural network with two layers of nodes (representing conjunctions and word orders, respectively) that can simulate the observed data pattern through inhibition-based competition of the alternative choices within the node layers.

1. Introduction

In modern Dutch, word order of subject, object(s) and finite verb in main clauses differs from that in subordinate clauses: verb-second (V2) and verb-final (VF), respectively. Explanatory clauses (expressing causative or argumentative relations) may manifest either order: verb-final is mandatory after the subordinating conjunctions *omdat* and *doordat* (both translatable as *because*), verb-second after the coordinating conjunction *want* ('since', 'for'). This difference mirrors the situation in German, with *weil* 'because, since, as' licensing verb-final, and *denn* 'for' licensing verb-second word order. In this paper, we focus on a phenomenon readily observable in spoken Dutch and German: verb-second word order in subordinate clauses introduced by *omdat/weil*. We will refer to the resulting structure as *omdat*-V2 and *weil*-V2. The clause in (1), from a recent paper on *omdat*-V2 by Degand (2016), is a case in point, with the finite verb *heb* in V2 position:

* Corresponding author.
E-mail addresses: gerard.kempen@mpi.nl (G. Kempen), harbusch@uni-koblenz.de (K. Harbusch).

https://doi.org/10.1016/j.langsci.2018.05.005
0388-0001/

(1) …*omdat ik heb tot half één tentamen*
... because I have till half one exam
'… because I have an exam till 12:30'

Following Reis (2013), we use the terms *explanatory clause* or *explanans* to refer to a clause whose proposition denotes an event or state of affairs causing or explaining another event or state—the effect. The clause describing the effect will be called *explanandum*. In line with the literature, we distinguish two types of explanantia (explanatory clauses, often called causative and argumentative, respectively). For convenience (presumably without loss of generality of the theoretical account developed below), we assume that explanandum clauses have the form of main (matrix, root) clauses, hence featuring verb-second position of the finite head verbs.

In the present paper, we describe the *omdat*-V2 phenomenon in more detail, comparing and contrasting it with *weil*-V2. The empirical material consists of frequency counts of various types of explanatory sentences extracted from the Corpus of Spoken Dutch (Corpus Gesproken Nederlands, CGN2.0; Hoekstra et al., 2001; van Eerten, 2007). We propose a theoretical account based on the sentence *production* processes underlying *omdat*-VF and *omdat*-V2 structures. This account extends the *weil*-V2 theory published recently by Kempen and Harbusch (2016). The key notion is inhibition-based competition between (1) subordinating and coordinating causative conjunctions, and (2) between verb-second and verb-final word order in the explanans clauses. These competitions are assumed to partly proceed independently and to sometimes miscommunicate.

We begin by summarizing the main properties of *omdat*/*weil*-V2 structures put forward in the literature (for recent surveys, see Persoon et al., 2010; Antomo and Steinbach, 2010; Reis, 2013; Abraham, 2016; Kempen and Harbusch, 2016). We take these to be uncontroversial.

1. *Omdat*/*weil* explanans clauses with noncanonical V2 mainly occur in spontaneously (extemporaneously) produced spoken utterances, and to a lesser extent in nonstandard written language (Schäfer and Sayatz, 2017). They are virtually absent in formal/edited written text.
2. They do not occur in subordinate explanans clauses that precede the explanandum (hence, *omdat*/*weil*–V2 clauses 'trail' behind their explanandum clause).
3. Explanans clauses introduced by *want*/*denn* always follow their explanandum (and are traditionally analyzed as the posterior member of a clausal coordination).
4. *Omdat*/*weil*-V2 explanans clauses do not occur as isolated explanantia, e.g., in elliptical answers to *waarom* 'why' questions.
5. There is no mirror image phenomenon consisting of *want/denn* clauses with VF order.

In the next section, we describe the design and the main results of the corpus study into the incidence of VF and V2 orders in causative/argumentative sentences of spoken Dutch. Our account of these data and of the similarities and differences between the Dutch and German data patterns is developed in Section 3. In Section 4, we outline how the theoretical account can be implemented in the form of a simple neural net, and how it deals with the (near-)absence of V2 after some low-frequent subordinating explanatory conjunctions in both target languages. Section 5 summarizes the theoretical account put forward in the two preceding sections. In final Section 6, we briefly discuss relations with extant linguistic theories about *omdat/weil*-V2.

2. The corpus study: design and results

The CGN2.0 treebank contains about 130,000 spoken sentences (dialogue turns) from varied content domains (news, telephone conversations, lectures, etc.). Not all of them were produced spontaneously; we discarded about 3800 sentences with read speech. The sentences had been annotated syntactically with relatively theory-neutral dependency graphs. Using TIGERsearch (König and Lezius, 2003) and JAVA programs written by ourselves, we extracted all sentences that contain one or more tokens of the conjunctions *omdat, want,* or *doordat*, with the explanandum preceding the conjunction. This yielded a set of 4220 explanandum-cum-explanante sentences, that is, sentences that include a trailing explanans clause with enough lexical material to determine verb placement unambiguously as either VF or V2.[1] The explanatory conjunctions we extracted were *want*, *omdat*, and *doordat*.[2] The resulting data are shown in Table 1.

[1] In both target languages, the amount of explanandum-cum-explanante sentences fluctuates around 3% of the sentences annotated in the form of a dependency graph: 3.3 in the Dutch, 2.6 in the German corpus.

[2] We had to relinquish our plan to analyze clauses introduced with *daar*. This subordinating conjunction occurred only once, followed by an incomplete clause.

In response to a suggestion by one of the reviewers, we also extracted sentences containing the conjunction *aangezien* 'as, seeing (that), given (that)'. We found 10 exemplars unambiguously classifiable w.r.t. word order. Of these, five were leading, four were trailing clauses, one was an isolated clause. In one leading and one trailing clause, *aangezien* constituted a conjunction together with the complementizer *dat* (*aangezien dat*). See also footnote 11.