Glossa a journal of general linguistics

Kempen, Gerard and Karin Harbusch. 2016. Verb-second word order after German *weil* 'because': Psycholinguistic theory from corpus-linguistic data. *Glossa: a journal of general linguistics* 1(1): 3. 1–32, DOI: http://dx.doi.org/10.5334/gjgl.46

RESEARCH

Verb-second word order after German *weil* 'because': Psycholinguistic theory from corpus-linguistic data

Gerard Kempen[1] and Karin Harbusch[2]

[1] Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands
Cognitive Psychology Unit, Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands
gerard.kempen@mpi.nl

[2] Department of Computer Science, University of Koblenz-Landau, Universitätsstraße 1, 56070 Koblenz, Germany

Corresponding author: Gerard Kempen

In present-day spoken German, subordinate clauses introduced by the connector *weil* 'because' occur with two orders of subject, finite verb, and object(s). In addition to *weil* clauses with verb-final word order ("VF"; standard in subordinate clauses) one often hears *weil* clauses with SVO, the standard order of main clauses ("verb-second", V2). The "*weil*-V2" phenomenon is restricted to sentences where the *weil* clause follows the main clause, and is virtually absent from formal (written, edited) German, occurring only in extemporaneous speech. Extant accounts of *weil*-V2 focus on the interpretation of *weil*-V2 clauses by the hearer, in particular on the type of discourse relation licensed by *weil*-V2 vs. *weil*-VF: causal/propositional or inferential/epistemic. Focusing instead on the production of *weil* clauses by the speaker, we examine a collection of about 1,000 sentences featuring a causal connector (*weil, da* or *denn*) after the main clause, all extracted from a corpus of spoken German dialogues and annotated with tags denoting major prosodic and syntactic boundaries, and various types of disfluencies (pauses, hesitations). Based on the observed frequency patterns and on known linguistic properties of the connectors, we propose that *weil*-V2 is caused by miscoordination between the mechanisms for lexical retrieval and grammatical encoding: Due to its high frequency, the lexical item *weil* is often selected prematurely, while the grammatical encoder is still working on the syntactic shape of the *weil* clause. *Weil*-V2 arises when pragmatic and processing factors drive the encoder to discontinue the current sentence, and to plan the clause following *weil* in the form of the main clause of an independent, new sentence. Thus, the speaker continues with a V2 clause, seemingly in violation of the VF constraint imposed by the preceding *weil*. We also explore implications of the model regarding the interpretation of sentences containing causal connectors.

Keywords: Verb-second word order; *weil* clause; epistemicity; grammatical encoding; lexicalization; corpus linguistics; German language

1 Introduction

In present-day spoken German, subordinate clauses introduced by the conjunction *weil* 'because' occur with two orders of subject, finite verb, and object(s). In addition to *weil* clauses with verb-final word order (henceforth "VF"; the standard order in subordinate clauses) one often hears *weil*-clauses with SVO, the standard order of main clauses ("verb-second", V2). The "*weil*-V2" phenomenon is restricted to colloquial language registers, virtually absent from formal (written, edited) German. Over the past decades, quite a few studies have proposed explanations for *weil*-V2 (e.g., Gaumann 1983; Günthner 1993; Wegener 1993, 1999; Keller 1995; Antomo & Steinbach 2010; Antomo 2012; Reis 2013). Many of these studies address observations suggesting that whereas *weil*-VF clauses express standard *propositional causality, weil*-V2 clauses express special types of discourse

relations called *epistemic causality* and *speech act justification* (see next section for examples of these three types). Another property believed to set *weil*-V2 apart form *weil*-VF concerns the level of *syntactic and prosodic integration* between *weil* clauses and the preceding main clauses: lower with V2 than with VF. Recently, Reis (2013) published a detailed account of these phenomena, which we introduce in the next section.

The impetus to the present study arose from earlier work (Kempen & Harbusch 2012, 2015) we carried out with the VERBMOBIL corpus: a large collection of spoken German dialogues annotated for prosodic and syntactic features (Wahlster 2000; see also Footnote 4). We examined the incidence of *weil*-V2 and *weil*-VF clauses in this corpus, expecting to find clues to the *grammatical encoding process* underlying the two constructions. Hence, our emphasis is on the production rather than the comprehension of the VF and V2 variants.

In Section 2, we summarize the linguistic account put forward by Reis (2013), and explain the terminology introduced above. In Section 3, we report the design and the results of our exploration of the VERBMOBIL corpus of spoken German. In Section 4, we develop a new theoretical account of *weil*-V2 phenomena partly prompted by the corpus findings. Finally, in Section 5, we consider theoretical issues and implications.

2 *Weil*-V2 according to Reis (2013): A summary

Reis's first illustration of a *weil*-V2 clause and its *weil*-VF counterpart is reproduced in (1a) and (1b), respectively. As for terminology, we will use the term "explanandum" to refer to (the meaning expressed by) the main clause preceding[1] the *weil* clause, and the term "explanans" to denote (the meaning expressed by) the ensuing *weil*-V2 or *weil*-VF clause itself.[2]

(1) a. Wir waren gestern schwimmen, weil das Wetter war so
we were yesterday swim because the weather was so
schön. (V2)
nice
'We were swimming yesterday because the weather was so nice.'

b. Wir waren gestern schwimmen, weil das Wetter so schön
we were yesterday swim because the weather so nice
war. (VF)
was
'We were swimming yesterday because the weather was so nice.'

Reis (p. 223) lists three formal properties of sentences that combine an explanandum and a *weil*-V2 explanans—properties she considers to be uncontroversial: (1) The two clauses are *prosodically unintegrated*; that is, they each have their own focus–background articulation. (2) The two clauses constitute a paratactic structure, implying that they are also *syntactically unintegrated*. (Properties (1) and (2) together constitute the above-mentioned prosodic/syntactic autonomy of the *weil*-V2 clause.) (3) The *weil*-V2 clause always occurs in *root clause position*, i.e., does not attach to embedded clauses. To these,

[1] It is generally recognized that *weil*-V2 phenomena never occur in *weil* clauses that precede the explanandum, i.e., are located in the forefield ("Vorfeld") of the latter. As regards the term forefield, we adopt (English translations of) the structural-linguistic descriptions of constituent order in main and subordinate clauses of German. One usually distinguishes five "topological" positions, called (from left to right): (1) *Vorfeld* (forefield; at most one constituent), (2) *Linke Satzklammer* (left sentence bracket, hosting the finite verb in main clauses; hence "verb-second", (3) *Mittelfeld* (midfield; zero or more constituents), (4) *Rechte Satzklammer* (right sentence bracket; hosting the finite verb of subordinate clauses ("verb-final") and non-finite verbs in main or subordinate clauses, and (5) *Nachfeld* (endfield; typically occupied by complement clauses, adverbial clauses, or extraposed constituents). See, e.g., Drach (1937) and Höhle (1986).

[2] We use the terms explanandum and explanans in two senses, one referring to the meaning underlying the clauses, the other one to the clauses themselves. We trust that this ambiguity is easily resolved in context.

Reis adds an—equally uncontroversial—interpretive property (p. 224): (4) The *weil*-V2 clause realizes an *autonomous speech act,* i.e., one that is independent from the speech act realized by the explanandum; and, by virtue of the semantics of *weil*, the explanans speech act has *assertional illocutionary* force.

The sentences in (2) illustrate the three types of causal interpretation licensed by explanantia that are integrated (2a) or unintegrated (2b–d) with their explananda (for details, see Antomo & Steinbach 2010: 16–19). Propositional explanations specify a causal relation between the event/state-of-affairs expressed by the explanans and the one expressed by the explanandum (2a-b). So-called *epistemic* explanations specify an *inference* process: The explanandum expresses the conclusion reached by the speaker, the explanans states the evidence/argument for the conclusion (2c). A special kind of epistemic explanation are speech act justifications: The explanans states the argument for the speech act the speaker has realized in the explanandum (in (2d): the assertion mentioned in the main clause). Propositional explanations can be realized with an integrated clause (*weil*-VF; as in (2a)) as well as an unintegrated clause (*weil*-V2; (2b)). Epistemic explanations and speech act justifications are realized with unintegrated *weil* clauses, where the lack of integration with the explanandum is evinced by V2 word order alone (2c), and/or by separate sentential prosodic contours for explanandum and explanans, as suggested by the punctuation in (2d).

(2) a. Mittags sind wir zurückgefahren, weil der Himmel ganz grau
at.noon have we returned because the sky very gray
war.
was
'In the afternoon we returned because the sky was very gray.'

b. Mittags sind wir zurückgefahren, weil der Himmel war ganz
at.noon have we returned because the sky was very
grau.
gray
(Antomo & Steinbach 2010: 17, ex. (22a,b))

c. Es hat geschneit, weil die Straße ist ganz weiß.
It has snowed because the street is totally white
'It has snowed because the street is totally white.' (Antomo & Steinbach 2010: 17, ex. (23a))

d. Ich habe den Job gekriegt. Weil das interessiert dich doch am
I have the job got because that interests you indeed the
meisten.
most
'I got the job. Because that is what you are primarily interested in, aren't you?' (Antomo & Steinbach 2010: 19, ex. (24a))

Contra Antomo & Steinbach, Reis argues that epistemic and speech act justifications do not necessarily require V2 word order in the *weil* clause. Syntactic and prosodic autonomy can license these interpretations (including assertional illocutionary force) even with VF word order. The examples in (3) illustrate this point with respect to an epistemic explanation (3a), and justifications of an interrogative (3b) and an imperative (3c) speech act.[3] Backslashes ("\") mark a major intonational break at the border between explanandum

[3] Reproduced from Reis (2013: 243), examples (28d), (29a), and (29b). Capitalized syllables indicate sentence accent. Example (3b) is originally from Antomo & Steinbach (2010: 19), example (24e).

and explanans, indicating that the *weil* clauses are prosodically unintegrated. (Stressed syllables are printed in small capitals.)

(3) a. Die BREMsen haben wohl versagt, (\) weil Sue an sich eine
the brakes have probably failed because Sue in fact a
VORsichtige Fahrerin ist.
careful driver is
'The brakes must have failed, (\) because in fact Sue is a careful driver.'

b. Bist du nervös? (\) weil du schon deine dritte Zigarette
are you nervous because you already your third cigarette
rauchst.
smoke
'Are you nervous? (\) Because you are already smoking your third cigarette.'

c. Räum bitte dein ZIMmer auf, (\) weil wir morgen nämlich
tidy please your room up because we tomorrow namely
BeSUCH kriegen.
visitors get
'Tidy up your room, please, (\) because, the fact is, we have visitors tomorrow.'

Why do syntactically/prosodically autonomous *weil* clauses afford a broader range of interpretations of the semantic/pragmatic relation between explanandum and explanans than integrated ones? (That is, autonomous *weil* clauses license all three types of explanation, non-autonomous ones license only the propositional type.) Here, Reis (2013: 248–250) adopts the account proposed and experimentally tested by Antomo & Steinbach (2010: 25ff). These authors argue that explanandum and explanans of *weil*-V2 sentences constitute two semantic/pragmatic units that are first processed separately, and whose discourse relationships are established in a second step. In contrast, *weil*-VF clauses and their explananda form a single semantic/pragmatic unit that can be processed in one step: "Less strongly connected structures seem to grant more interpretive leeway than more strongly connected structures" (Antomo & Steinbach 2010: 25; see also Scheutz 2001).

Figure 1 may serve to highlight the essentials of Reis's (2013) account of *weil*-V2 phenomena. For the empirical evidence, we refer to her paper and, insofar as the evidence

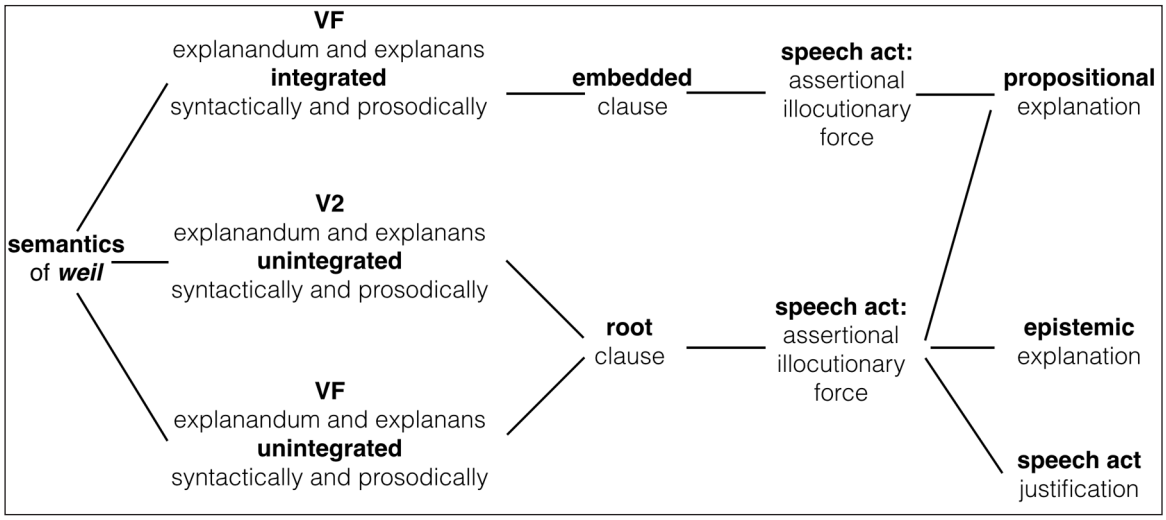

Figure 1: Essential aspects of the hearer-centered *weil*-V2 account proposed by Reis. The diagram specifies how combinations of conjunction, level of integration and word order, mediated via clause type and speech act with illocutionary force, predict explanation type.

is undisputed, to Antomo & Steinbach (2010). The schema brings out clearly that the account takes the perspective of the *hearer*: The combination of lexical, syntactic and prosodic properties of a sentence at the left-hand side serves as input, and a set of interpretive options emerges as output at the right-hand side.

The main goal of the present paper is to explore, by means of a detailed corpus study, how speakers grammatically encode causality under varying sentence production conditions. We not only take word order into account but also the selection of two other common causative connectors: *denn* and *da*. We expect that this will also yield insight into the production of the three types of explanations distinguished here.

3 When do speakers produce verb-second word order in causal clauses?

This section describes the design and the results of an exploratory corpus linguistic study into *language production factors* determining when and why speakers use V2 rather than VF word order during the spontaneous oral production of *weil* clauses.

3.1 The VERBMOBIL corpus

As data source we used the transliteration files of all German dialogues in the VERBMOBIL corpus, recorded during spoken conversations between two native speakers of German. Participants were over 1000 adult native speakers of German recruited from regions around Munich, Bonn and Kiel. The conversations were recorded between 1993 and 2000 as part of a large computational-linguistic project on spoken language translation. During simulated phone conversations, the speakers discussed arrangements for a joint trip to a meeting. Seated in different soundproof rooms, they could hear each other via an audio connection. In most cases, they could also see each other through a window between the rooms.[4]

The dialogue contributions had previously been transliterated in detail, as part of the VERBMOBIL project. From the transcripts of the dialogues, we extracted all turns containing one or more tokens of an adverbial clause that is preceded by the main clause, and is introduced by the causal connectors *weil*, *da*, or *denn* (i.e., we included all "trailing" causal clauses and disregarded "leading" ones).[5] We included *da* and *denn* because they

[4] The VERBMOBIL corpus (Wahlster 2000) consists of two parts. The VERBMOBIL1 data were collected in the period 1993–1996, comprising 1422 recordings by 885 speakers (https://www.phonetik.uni-muenchen.de/Bas/BasVM1.1README). The VERBMOBIL2 corpus was collected in 1997–2000. It comprises 810 recordings by 445 speakers. (These data also include English and Japanese conversations with translations. We have ignored these, of course.) The VERBMOBIL1 format was rewritten according to the new format guide lines for VERBMOBIL2 (cf. Burger et al. 2000). The task the speakers had to execute during the conversations has been described by Hess et al. (1995) as follows: "The Verbmobil domain is negotiation, and the task to be solved by speakers is to arrange meetings and plan a trip. Speakers are given a sheet of instructions consisting of a three-month and a week calendar with blocks indicating time slots that cannot be used; they are asked to introduce themselves, use a formal style of speech, and to arrange one to eight meetings of different durations (ranging from a short meeting to a four day business trip)."

[5] As for the number of leading causal clauses, the corpus contains 111 *da* clause (all VF) preceding the main clause, that is, a few more than the number of trailing *da* clauses (104). Leading *weil* clauses hardly occur: We found no more than 5 exemplars (all VF). However, most of them are dubious cases since they were not only followed but also preceded by a main clause, without prosodic tags (full stops) demarcating sentence boundaries. The only clear case is the following (punctuation as in the transcription):

(i) ja, weil es bei mir am Freitag sehr schlecht aussieht, da bin ich den ganzen Tag
yes because it with me on Friday very bad looks then am I the whole day
über beschäftigt, würde mir der Sonntag ziemlich gut passen. da hab' ich
during busy would me the Sunday rather well fit then have I
keinerlei Verpflichtungen. Sonntag, der fünfundzwanzigste, wäre das in
none-at-all obligations Sunday the 25th would-be that in
Ordnung?
order
'Yes, because it looks bad for me on Friday, then I'm busy (during) the whole day, would Sunday suit me rather well. Then I have no obligations at all. Sunday the 25th, would that be OK?

also express causality but do not allow any verb placement freedom: *da* always introduces a VF clause, *denn* always a V2 clause. We found 721 usable clauses introduced with *weil*. Of these, 385 have VF word order (53%), and 336 occur with V2 (47%). There are 160 sentences with a *denn*, and 104 with a *da* clause. These numbers (Figure 2) do not include a small number of sentences that we had deemed unusable: Some clauses could not be classified unambiguously as V2 or VF because, apart from the subject, all nonverbal constituents were located in the endfield (e.g., as a complement or adverbial clause), or because they were interrupted and aborted before a potential V2 position of the verb. The resulting collection consisting of 985 usable sentences was subjected to the analyses we report below.[6]

To convey an impression of similarities and differences between the extracted clauses, we list one sample sentence from each of the four classes; the four explananda are very similar in that they all designate reasons why a proposed time slot does not suit the speaker (compare (4a–d); connectors and finite verbs in the causal clauses are highlighted in bold font).

(4) a. oh, das ist schlecht, **weil** die Wochenenden bei mir so ziemlich
oh, that is bad because the weekends for me so rather
ausgebucht **sind**. (*weil*-VF)
booked-up are
'Oh, that's bad, because my weekends are pretty much booked up.'

b. das ist schlecht, **weil** da **hab'** ich einiges vor. (*weil*-V2)
that is bad because then have I a.few.things ahead
'That's bad, because then I have already planned a few things.'

c. oh, das is' bei mir schlecht, **denn** da **bin** ich vom 26sten bis
oh that is for me bad for then am I from.the 26th till
27sten in Zell. (*denn*-V2)
27th in Zell
'Oh, that suits me badly, for then I am in Zell from 26th till 27th.'

d. das ist schlecht, **da** ich um zwölf Uhr ein Arbeitstreffen hier
that is bad since I at twelve o'clock a work-meeting here
in München **habe**. (*da*-VF)
in Munich have
'That's bad, because at 12 o'clock I already have a work meeting here in Munich.'

The transcriptions do not only contain the words said by the speakers but also tags denoting syntactic and/or prosodic features. Given the importance of the level of integration of explanandum and explanans clauses in the *weil*-V2 literature (see Section 2), we focused on two types of tags immediately preceding, during, or immediately following the connector: (1) those marking syntactic/prosodic clause boundaries, and (2) those denoting filled or unfilled speech pauses. The former consist of commas, full stops, and question marks; the latter mark disfluencies: hesitations, repairs, and editing terms. We have treated full stops and question marks as signs of relatively low levels of integration. Commas not accompanied by signs of disfluency were taken to reflect a relatively high level of integration. (The annotators had been instructed to insert a comma at the boundary between two clauses.) Combinations of a comma and a disfluency tag were assumed

The main clause *da bin ich den ganzen Tag über beschäftigt* 'then I'm busy the whole day' is an interjection; subject-verb inversion in *würde mir der Sonntag ziemlich gut passen* 'would Sunday suit me rather well' is not due to stylistic inversion but demonstrates that the *weil* clause occupies the forefield of the main clause (*würde . . . passen*); see also note 1.

[6] An Excel file with the complete collection can be obtained from the second author upon request. See also Appendix A.1 for some varied *weil*-V2 examples.

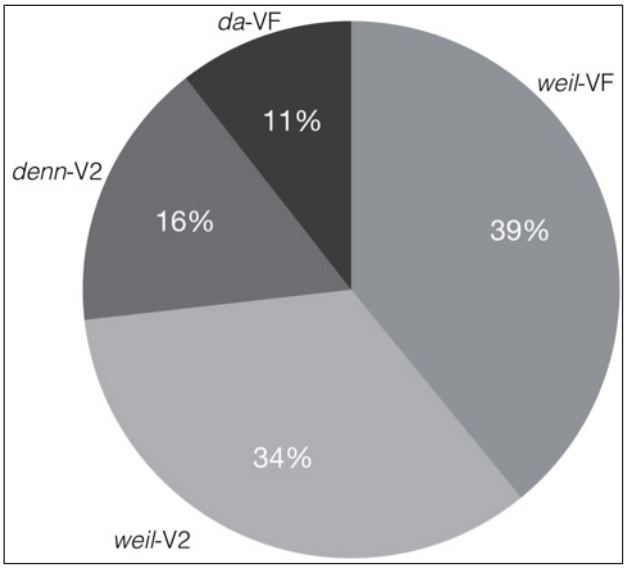

Figure 2: Distribution of four classes of causal clauses in the VERBMOBIL corpus.

to represent intermediate integration. See Section 3.3 for details, and Appendix A.1 for annotated examples.[7]

3.2 Defining and recognizing explanation types

Our attempts at classifying the causal clauses into the types distinguished in the literature (propositional explanations, epistemic explanations, and speech act justifications) turned out to be harder than anticipated. Here are the definitions we tried out.

With respect to epistemic causality, we assume it involves *Forward Causal Inference* (*FCI*) or *Backward Causal Inference* (*BCI*; also known as *abduction*[8]). Both types of inference include a *causal chain* (*CC*),

Causation(Effect:Prop2, Cause:Prop1),

where *Prop1* is a proposition denoting some observed event or state-of-affairs causing another event or state-of-affairs denoted by proposition *Prop2*.

FCI can now be characterized informally as follows. In FCI, the speaker claims *Prop2* to be true (justified) based on a two-part argument consisting of (1) evidence that *Prop1* is true, and (2) knowledge of the plausibility/likelihood of the event denoted by *Prop1* being the cause of the event denoted by *Prop2*:

Justification(Claim:Prop2,
Arg:[Prop1 & Probable(Causation(Effect:Prop2, Cause:Prop1))])

[7] For details, we refer to the corpus documentation (in German), available from http://www.bas.uni-muenchen.de/forschung/Verbmobil/VMtrlex2d.html. In Appendix A.2, we also present English translations of the most important tags and tagging rules. Notice also that, for the calculations reported in the present paper, we ignored tags not clearly related to cognitive processes in the speaker, e.g., coughing, breathing, and sounds produced by the recording equipment. That is, we only counted tags referring to hesitations (filled and unfilled speech pauses).

[8] A German term for abduction is *reduktiver Schluss* 'reductive inference'. For background information on abduction, see Hobbs et al. (1993). Pasch et al. (2003: 395–397) also used abduction in definitions of epistemic explanation types.

We realize the adjectives *backward* and *forward* in the definitions of FCI and BCI hereafter may be confusing because "forward inference", i.e., from effect to cause, means going backward in event time. However, we trust that this disadvantage is outweighed by the advantage of correspondence between direction of inferencing and order of mention in explanandum-cum-explanante sentences.

For example, based on knowledge that Hans was speeding (*Prop1*), and that speeding drivers are often ticketed (*Prop1 & (Causation(Effect:Prop2, Cause:Prop1)))*, one can rightly claim that *Hans got a ticket because/for he was speeding*.

In BCI (abduction), the roles of evidence and claim played by the two propositions are reversed, but not those of cause and effect:

Justification(Claim:Prop1,
Arg:[Prop2 & Probable(Causation(Effect:Prop2, Cause:Prop1))])

A person saying *Hans was speeding because/for he got a ticket* deploys this inference pattern to justify his/her claim about the likelihood of Hans having been speeding. From evidence that Hans was ticketed (*Prop2*), the speaker reasons backward—using the same CC as in the previous example—and makes the claim that Hans was speeding (*Prop1*).

Interrogative and imperative speech act justifications (see examples (3b) and (3c)) are similar to BCI, the only difference being that "*Claim:Prop1*" is replaced by "*Question:Prop1*" and "*Command:Prop1*", respectively. Propositional explanations are simpler than any of the foregoing types: They consist only of a causal chain, whereas a CC is a proper part of any explanation based on logical inference. Another way of characterizing the difference between propositional (CC) and epistemic (FCI, BCI) explanations is as follows. In a propositional explanation, the set of properties attributed to the event that is described by the explanandum, receives an additional member (the property denoted by the explanans proposition: its cause); i.e., the explanandum event is "modified" by the explanans. In an epistemic explanation, the inference act justifying the conclusion/claim expressed in the explanandum proposition receives an additional specification: its arguments. Hence, the explanans does not "modify" the event denoted by the explanandum but the inference act yielding the explanandum proposition.

When inspecting the corpus sentences, we did not find *weil*, *denn*, or *da* clauses unambiguously classifiable as epistemic (as opposed to propositional explanations) by the above definitions.[9] Virtually all the sentences express some form of psychological (motivational), social or institutional causality, but it was not possible to reliably distinguish cases expressing "mere" causality (CC) from cases where causality functions within FCI or BCI. For instance, it is hard to determine whether a person who produces the two clauses in *Hans got a ticket because he was speeding* in a prosodically more or less integrated fashion, asserts a causal relationship (CC), or lays a claim based on causal inference (FCI). Corpus sentences such as those in (4) and in Appendix A.1 testify to the problem. Hence, we will be unable to present numerical data pertaining to explanation types—except for 26 speech act justifications, which are easily recognizable by their explananda (interrogative main clauses tagged with a question mark: 22 cases with *weil*-V2; 2 with *weil*-VF; 2 with *denn*; 0 with *da*).[10] This means the corpus is unsuited to study the effect of epistemicity, in particular BCI, on the choice of V2 vs. VF word order in *weil* clauses.

However, the corpus contains similar proportions of *weil*-VF and *weil*-V2 clauses. In combination with the annotations added to *weil*, *da* and *denn* sentences, this enables us

[9] This agrees with earlier corpus studies (Scheutz 2001; Dittmar & Bressem 2005: 112–114) insofar as, also in those corpora, epistemic and speech act justifying explanantia are rare.

[10] Here is one example of an interrogative speech act justification with *weil*-VF (cf. example (3b) in Section 2):

(i) ach so, ich überlege gerade, schaffen wir des dann überhaupt? weil wir ja
oh well I think just manage we that then at.all because we indeed
an'erthalb Arbeitstage einplan' müssen für unser'n Auf'n'halt dort, 'n unsrer Filiale.
one.and.a.half work.days plan have.to for our stay there in our branch
'I see; I'm just thinking, can we manage that at all? Because we have to plan a one-and-half-day stay there, in our branch.'

to explore verb placement determinants other than explanation type (if any). Identifying such factors thus becomes the main goal of the present study. To this purpose, we focus on corpus data tapping into how the speakers delivered the transitions between explanandum and explanans clause, and how they shaped the explanans clauses themselves.

3.3 Main quantitative results

After initial explorations of the final collection of sentences, we decided to rate the explanandum-to-explanans transitions for prosodic integration of its delivery according to the following criterion. Transitions marked only by a comma were assigned the highest level of integration. If such a transition was additionally tagged with one or more disfluency markers, its integration level was coded as intermediate. The lowest level of integration was assigned to transitions tagged with a full stop or a question mark. In view of the sparsity of the latter type of transitions, we did not subdivide them into cases with vs. without disfluency tags. Figure 3 shows how the three levels of prosodic integration are distributed over clause types.

Initial inspection of the sentence collection suggested that constituent order within the explanans clause is correlated with the syntactic shape and position of the subject NP of the clause.[11] In the large majority of VF clauses, the conjunction is immediately followed by a personal-pronoun subject (*ich* 'I', *du* 'you', *er* 'he', *sie* 'she', *es* 'it'; *wir* 'we', *ihr* 'you', *Sie* 'you', *sie* 'they'), whereas this is the case in only a minority of the V2 clauses. Hence, we also report counts of personal-pronoun subjects and the position of these subjects in the clauses. If a clause contains a personal-pronoun (PPro) subject, we distinguish between

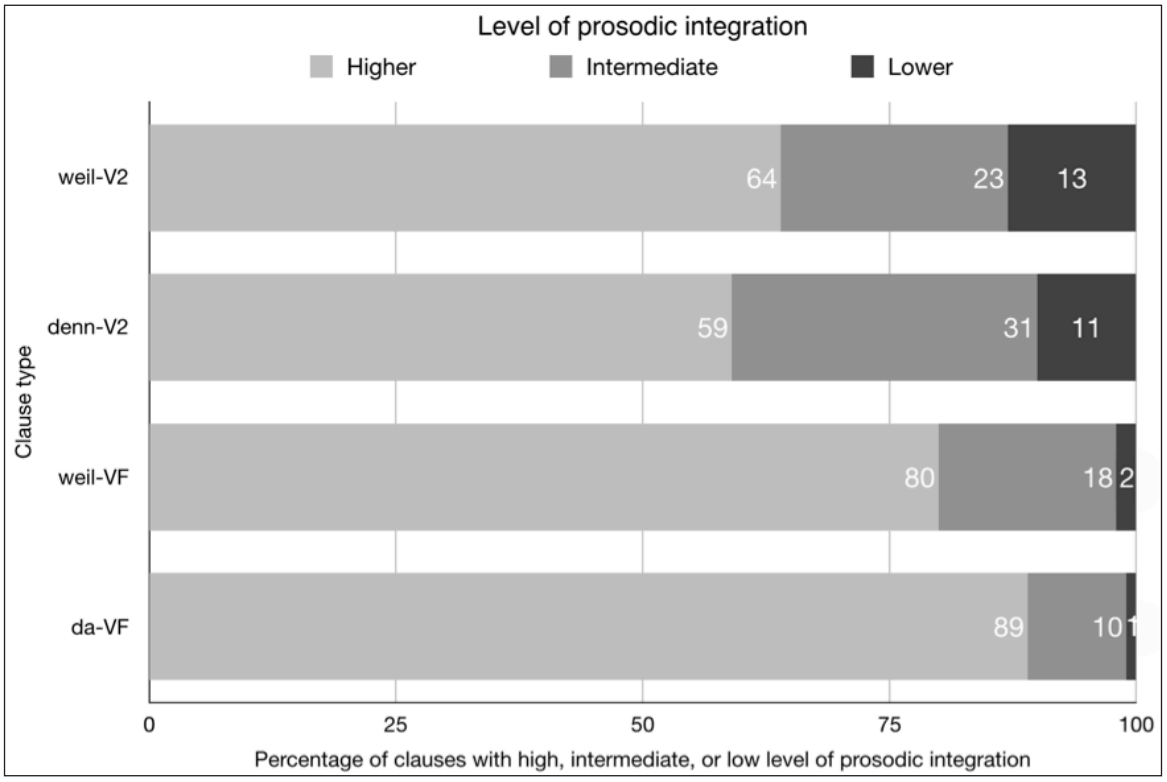

Figure 3: Distribution of integration levels over the four types of explanans clauses. (Percentages based on the raw scores in Table 1.)

[11] For a similar observation in a corpus of causal sentences produced by children speaking Swiss German, see Schönenberger (2000, 2010).

placement immediately after the conjunction (at what we informally call the *Wackernagel* position; label: PPro + W +), or at a later position (PPro + W–); subject NPs with any other type of head word (including demonstrative or indefinite pronouns) are coded as PPro–. See Figure 4 for the distribution over clause types. Note that, in V2 clauses, subjects carrying the "PPro + W–" tag actually occur in the midfield, due to subject-verb inversion. In VF clauses, hardly anything can be placed in between the conjunction and the personal-pronoun subject; we found only two[12] cases (one being *weil an anderen Sonntagen ich keine Zeit habe* 'because on other Sundays I have no time'). This suggests that, if speakers—for information-structural reasons[13] or otherwise—wish to put a non-subject constituent before a personal-pronoun subject, they better build a V2 clause. In what follows, we treat the occurrence of PPro + W + subjects in explanans clauses as a factor promoting VF word order because these subjects facilitate word order decisions in VF clauses (see Section 4.2).

Table 1 shows how many tokens of the various combinations of properties we found in the final sentence collection. The core of these data is presented graphically in Figure 5.

Also noteworthy is the distribution of disfluencies adjacent to the conjunction. Before the conjunction (i.e., at the end of the explanandum), the four clause types attract the same number of disfluencies (around 12%; see Figure 6). After the conjunction, the V2 explanantia are accompanied by somewhat more disfluencies than their VF counterparts.

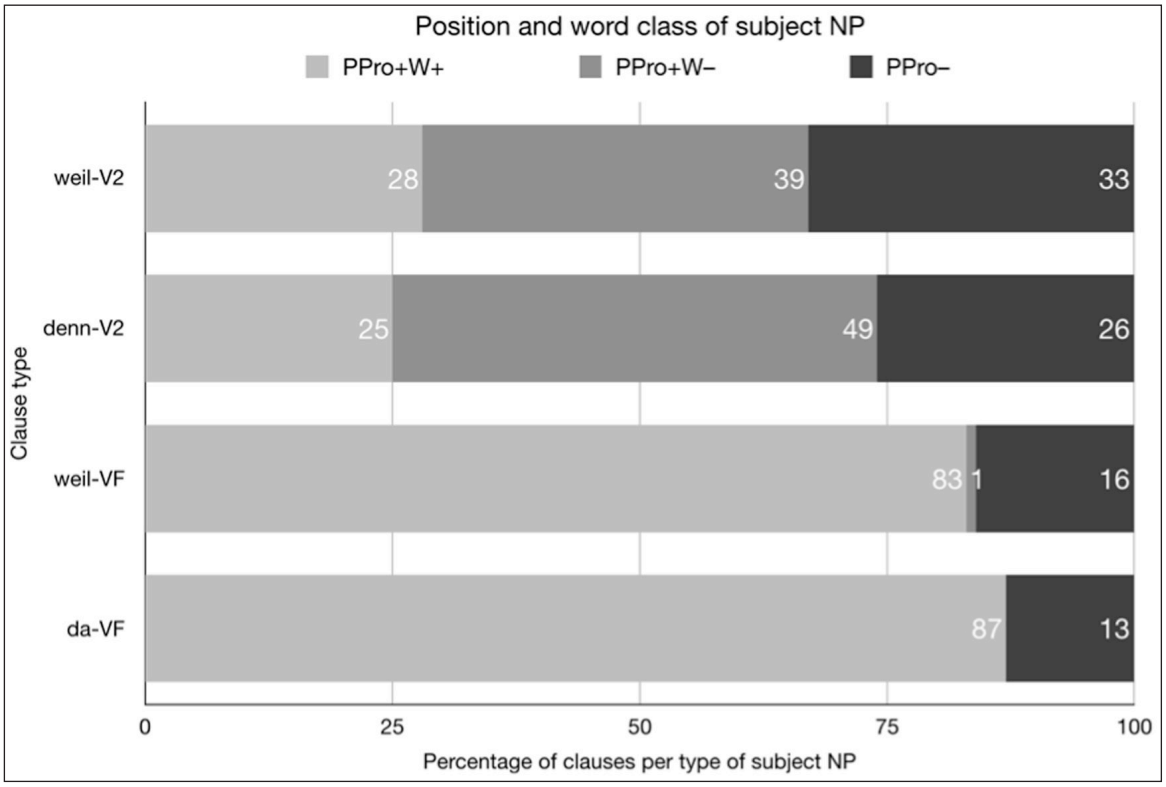

Figure 4: Distribution of subject NP types over explanans clause types. For abbreviations, see text. The percentages are based on the raw scores in Table 1.

[12] Not counting a third case (*weil sich's um ein fünftägiges Arbeitstreffen in der Filiale AGTR in Bonn handelt* 'because it has to do with a five-day work meeting in the AGTR branch in Bonn') where a reflexive pronoun (*sich* 'itself') precedes the cliticized personal-pronoun subject '*s* (= *es* 'it'). This sentence is grammatically well-formed. The early position of *sich* need not reflect the influence of an information-structural factor.

[13] Wegener (1993: 302) notes that *weil*-V2 extends the speaker's expressive means by allowing Topicalization and Left Dislocation—two other effects related to information structure. Example (9) in Appendix A.1 illustrates Left Dislocation.

Higher level of prosodic integration

Clause type	PPro+ W+	PPro+ W–	PPro-	Total
weil-V2	56	92	67	215
denn-V2	23	54	17	94
weil-VF	267	2	39	308
da-VF	82	0	11	93
Total	428	148	134	710

Intermediate level of prosodic integration

Clause type	PPro+ W+	PPro+ W–	PPro-	Total
weil-V2	27	24	27	78
denn-V2	10	20	19	49
weil-VF	47	0	21	68
da-VF	7	0	3	10
Total	91	44	70	205

Lower level of prosodic integration

Clause type	PPro+ W+	PPro+ W–	PPro-	Total
weil-V2	10	16	17	43
denn-V2	7	5	5	17
weil-VF	7	1	1	9
da-VF	1	0	0	1
Total	25	22	23	70

Totals for different connectors

Clause type	Total
weil-V2	336
denn-V2	160
weil-VF	385
da-VF	104
Total	985

Table 1: Numbers of causal clauses of types defined in the present section.

That the delivery of *weil*-V2 clauses tends to proceed less fluently than that of *weil*-VF clauses is also reflected by decelerated pronunciation of the conjunction itself (marked, in the transcriptions, by a delay tag before the final consonant of *weil*): These slow-downs occurred 12 times in total, but most of them (10) appeared in clauses that developed into *weil*-V2s.

Here are summaries and brief discussions of the results reported above:

(#i) A substantial proportion (46.6%) of *weil* clauses adopts V2 word order (Section 3.1). This percentage resembles the one obtained by Dittmar & Bressem (2005) in a corpus of conversations recorded in Berlin around 1995, but deviates substantially

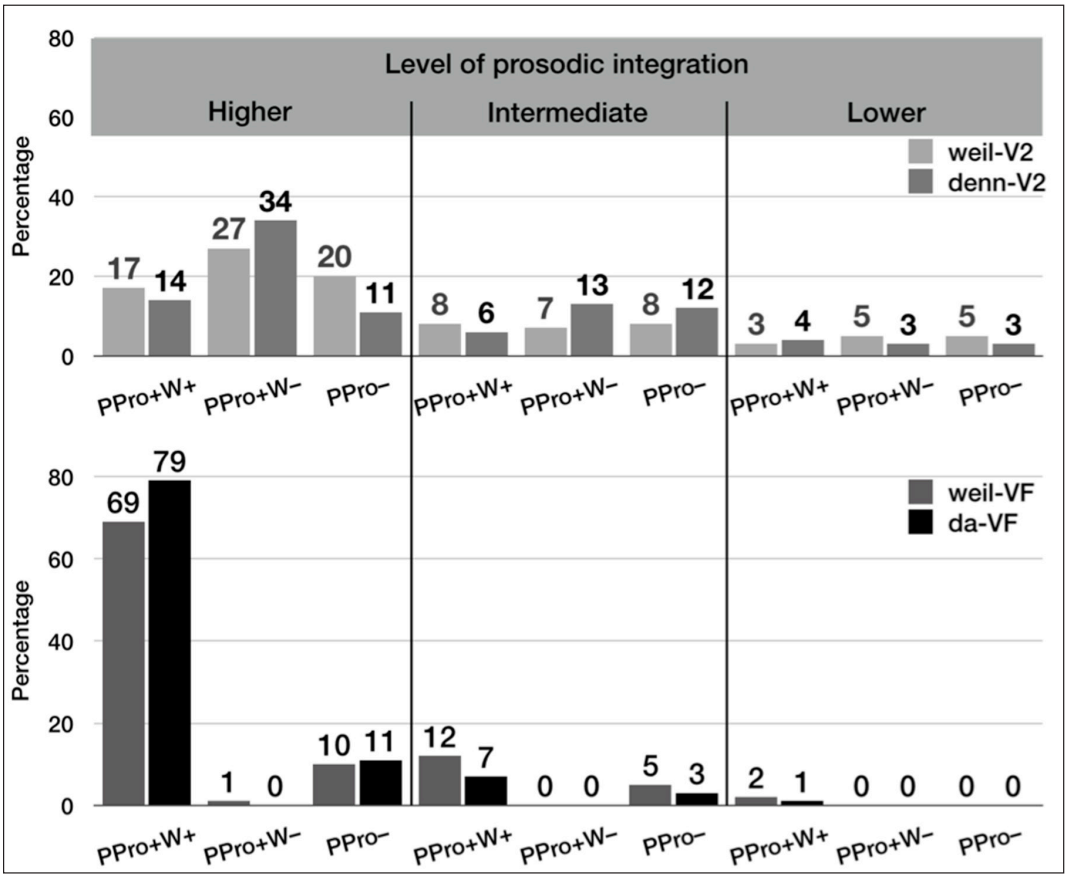

Figure 5: Graphical representation of the numbers in Table 1 expressed as percentages of the totals in the lower panel of the Table. Top: V2 clauses. Bottom: VF clauses. Abbreviations are explained in the text.

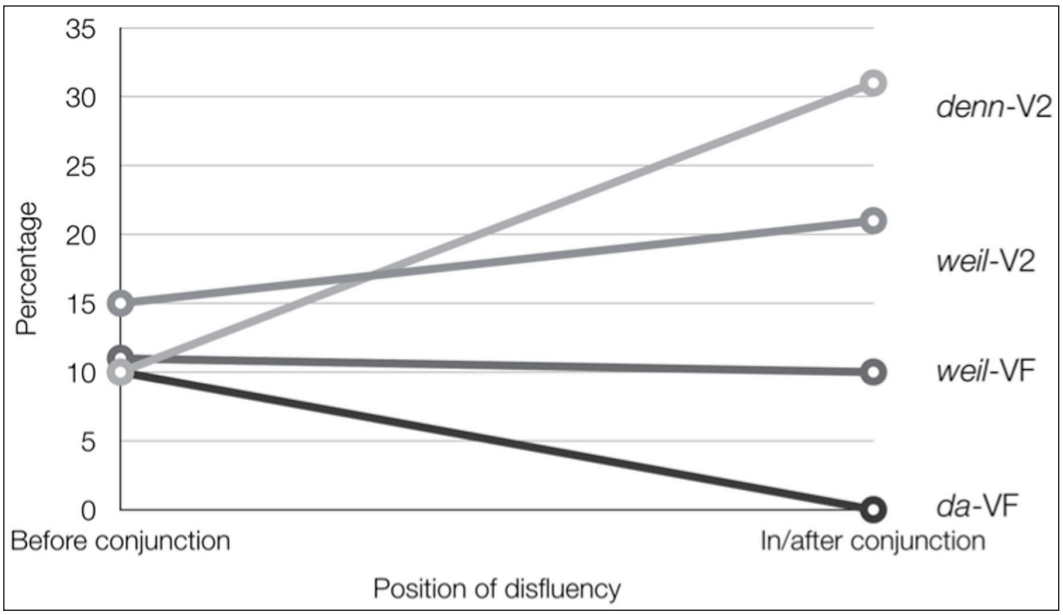

Figure 6: Breakdown of disfluencies according to their position immediately before, in, or after the conjunction. Note: The percentage of in/after-conjunction disfluencies for weil-V2 clauses (21%) includes 17 cases where the annotators inserted a comma tag after weil. As these commas are more likely to represent a disfluency than a syntactic transition, we included them in the graph. Without them, the percentage drops to 16. For denn, there are 4 similar cases; without them, the percentage drops to 29.

from the data reported by Freywald (2010) for spoken corpora recorded between 1955 and 1974: only 8.3%. We will not go into the reasons for this cross-corpus variability, which may be due to differences between participating speakers, regions, speech situations, and/or to historical developments (see, e.g., Freywald (2010) for data and references to the literature).

(#ii) However, no *weil*-V2 clauses we registered could be classified unambiguously as expressing an epistemic explanation in the abduction-based definition put forward in Section 3.2 (BCI).

(#iii) In the majority of cases speakers accomplish the transition from the leading explanandum clause to the trailing explanans clause smoothly, without overt disfluencies (see Figure 3, and panel "Higher level of prosodic integration" of Figure 5). For VF clauses, this is the expected pattern. But from the *weil*-V2 literature one might get the impression that, more often then not, clauses are delivered with speech pauses before or after the conjunction. However, this is not borne out by the data in our corpus, where 64% of the *weil*-V2s is delivered without disfluencies (215 out of 336; Table 1). The corresponding proportion of *denn*-V2s is a little lower but still the majority (59 percent: 94 out of 160 cases). This data pattern amounts to a partial dissociation between prosodic and syntactic integration insofar as a relatively high level of prosodic integration appears to go together with a relatively low level of syntactic integration (V2 < VF). We also observed some instances of the opposite mismatch: cases of intermediate or lower prosodic integration marked by a relatively high level of syntactic integration (19% *weil*- and 11% *da*-VF cases in Figure 5, middle and right columns; examples from the literature are (3a) and (3b) mentioned in Section 2). The possibility of rather high levels of prosodic integration demonstrated by *weil*- and *denn*-V2s, and low such levels by *weil*- and *da*-VFs, received little attention in the literature, which focused on epistemic explanations (abductions) and speech act justifications, i.e., on *weil*-V2 structures where prosodic integration levels indeed tend to be low.

(#iv) Immediately before the conjunction, disfluencies are relatively rare in V2 as well as in VF clauses. Immediately after the conjunction (and even during the pronunciation of the conjunction), disfluencies increase in V2 but not in VF clauses (Figure 6).

(#v) In the large majority of VF clauses, the subject role is fulfilled by an NP consisting of a personal pronoun occupying the so-called Wackernagel position, i.e., adjacent to the conjunction. The corresponding percentages in V2 clauses constitute the minority (Figure 4).

(#vi) The two types of VF clauses behave remarkably similarly with respect to the distribution of higher, intermediate and lower levels of integration, and the position and word class of the head word of subject NPs. The data pattern for the two V2 clause types resemble one another as well, but differ considerably from the VF pattern.

In the next section, we develop a production-based hypothesis about the origin of V2 word order in *weil* clauses based on these data. We seek this origin in special properties of the *lexicalization* and *grammatical encoding processes* in the speaker who is about to produce a *weil*-V2 clause.

4 The sentence production process underlying *weil*-V2 clauses

The goal of this section is to account for the following three general findings reported above:

(i) the partial dissociation between prosodic and syntactic integration of explanandum and explanans (many instances of relatively highly integrated V2 explanantia; finding (#iii) above);

(ii) the fact that the VF vs. V2 choice (findings (#iv) and (#v)) is influenced by the size of the cognitive processing capacity needed to grammatically encode the explanans; and
(iii) the strong similarity in both these respects between *weil*-V2 and *denn* clauses on the one hand, and between *weil*-VF and *da* clauses on the other (finding (#vi)).

Before turning to the model, we need to make explicit some important assumptions concerning the three causal connectors we are focusing.

4.1 Some syntactic and pragmatic properties of weil, denn, *and* da

It is generally recognized that *denn* is a paratactic rather than a coordinating conjunction (Höhle 1986; Pasch et al. 2003; Freywald 2009; Breindl 2009; Reis 2013: 225). Important arguments are the following. Unlike coordinating conjunctions such as *und* 'and' and *aber* 'but', *denn* imposes V2 word order on the clause it introduces (5a). *Denn* explanans clauses cannot get conjoined with VF explananda: In (5b), the *denn* clause can only be interpreted as paratactically conjoined with the *erwarte* 'expect' clause. Furthermore, *denn* does not license coordinate ellipsis (Antomo & Steinbach 2010: 27; e.g. (5c)); and two or more *denn* clauses cannot be coordinated (Breindl 2009: 280; compare examples (5e/f)). Finally, as is frequently noted in the literature, *denn* clauses cannot be within the scope of negation and quantifiers in the explanandum. (This also applies to the English causal connector *for*: The sentence *John didn't leave for he was ill* entails that John stayed; it cannot mean that John left, although for a reason other than being ill.)

(5) a. *Ich erwarte, dass du heute zuhause bleibst, denn du krank bist.
I expect that you today home stay for you ill are
'I expect you to stay home today because you are ill.'
b. Ich erwarte, dass du heute zuhause bleibst, denn du bist krank.
c. *Du bleibst heute zuhause, denn ~~du~~ bist krank.
d. Du bleibst heute zuhause und ~~du~~ bist krank.
e. Du bleibst zuhause, denn du bist krank und du könntest uns anstecken.
f. *Du bleibst zuhause, denn du bist krank und denn du könntest uns anstecken.
you stay home for you are ill and for you could us infect
'You stay home for you are ill and you could infect us.'

These properties of *denn* can be summarized in a simple rule: *Denn* is obligatorily followed by a syntactically entirely independent (unintegrated, autonomous) sentence. No special constraints seem to apply to the grammatical shape of this sentence, which can be of any type (declarative, interrogative, or imperative; simple, complex, or compound). Hence, after *denn* has delegated to an independent sentence the task of grammatically encoding the explanans content, this task can be carried out in the same manner as the task of grammatically encoding the content of an arbitrary subsequent sentence *not* preceded by *denn* (that is, only the usual constraints between consecutive sentences in connected discourse hold). But, although the *denn* explanans is syntactically independent of its explanandum, the rhetorical/discourse relation expressed by the connector still holds between the connected propositions.[14]

This excursion into *denn* enables us to make the crucial point that the mentioned *denn* properties hold for *weil*-V2 as well. The reader can verify this for her/himself by substituting *weil* for *denn* in examples (5b/c/f). It follows that, if *weil* precedes a V2 construction,

[14] The notations "*denn*-V2" and "*weil*-V2" actually are infelicitous insofar as suggesting that the V2 property of the clause following the connector is licensed by the connector itself. However, on the proposed account, V2 is imposed by general word order rules for main clauses.

this construction is best viewed as a grammatically fully independent sentence as well. (This obviously can only be true for *weil* explanantia that, like *denn* clauses, assert a proposition beyond the scope of explanandum constituents.)

If the explanans following the connector *denn* or *weil* is indeed encoded as a new sentence rather than as a clause licensed by *denn* or *weil*, the conceptualization process engendering the explanandum proposition must have terminated prior to embarking on the explanans—similar to the course of events during the production of any two consecutive sentences in a discourse. The resulting proposition is "frozen" and will not be modified anymore (except in the course of self-corrections due to "changes of mind"). We will refer to this "wrap-up" as *closure* of the explanandum. Closure of a proposition immunizes it against amendments. However, a closed proposition can still be embedded wholesale into an overarching proposition. Specifically, an explanandum proposition can become an argument in a causal chain or in an inference act (see Section 3.2). Importantly, closure of an explanandum proposition does not necessarily lead to closure of the *syntactic* structure encoding this proposition; this structure can still be extended, e.g. with a subordinate clause encoding the cause of an overarching (superordinate) causal chain.

If *denn/weil-V2* explanantia indeed constitute new sentences and their production is preceded by pre-explanans closure of the explanandum proposition, then this implies that the explanandum and explanans contents are conceptualized in two steps, i.e., *incrementally*,[15] each step engendering a proposition possessing (assertional) illocutionary force. However, not all explanandum-cum-explanante sentences are produced in this manner. Consider the examples in (6).[16] Sentence (6a), with *weil*-VF, can be interpreted as resulting from *integrative* (i.e., non-incremental) conceptualization, yielding a single proposition with assertional illocutionary force: Some drivers in the domain of discourse owe their tickets to speeding. Within this assertion, the explanans expresses *new* information, whereas the explanandum encodes conceptual content that is *presupposed* (given, old). A speaker who intends to convey the explanans proposition as new information, will *foreground* (i.e., assert) it. Propositions that are already shared with the interlocutor—technically called *presuppositions*—, will remain *backgrounded* because they already belong to the interlocutor's knowledge base. Mentioning a presupposed proposition serves as a signal to the addressee that the proposition should be activated, i.e., retrieved it from long-term memory and placed in her/his working memory (Dryer 1996).

(6) a. Einige Fahrer kriegen regelmäßig ein Knöllchen, weil sie
some drivers get regularly a ticket because they
immer rasen. (VF)
always speed
'These drivers get tickets regularly because they are always speeding.'

b. Einige Fahrer kriegen regelmäßig ein Knöllchen, weil sie
some drivers get regularly a ticket because they
rasen immer. (V2)
speed always

Weil-V2 example (6b), on the other hand, expresses two propositions/assertions brought about by an incremental, two-step conceptualization process. This mode of conceptualization tends to be reflected prosodically by two separate sentential intonation contours,

[15] See Bock et al. (2003), Lee et al. (2013), and van de Velde et al. (2015) for theoretical and empirical discussions of incremental and non-incremental sentence production.

[16] The commas before the conjunctions in (6a–b) are due to a German spelling rule and do not reflect interpretation or prosody.

each spanning one clause (but see Section 4.3). Integrative conceptualization is more likely to yield a single sentential intonation contour spanning both clauses.

We now briefly note some properties of *da*, the third causal connector. *Da* cannot be within the scope of a negation element in the explanandum (see Sentence (7)) and often introduces epistemic explanations (Pasch 1983, 1997; Pasch et al. 2003: 397–399; see also Stede 2008: 224–225 and Bisiada 2013: 14). In this respect, *da* resembles *denn*. However, whereas *denn* introduces an explanans in the form of a main clause with foregrounded information, *da* expresses old/presupposed, hence backgrounded information. *Da* indeed does not start up a new sentence but forms a subordinate clause within the same sentence as the explanandum.

(7) Ich hab' noch nie ein Knöllchen gekriegt, weil/*da ich gerast
I have yet never a ticket got because I speeding
bin. (VF)
am
'I never got a ticket because I was speeding.'

In contrast, *weil* is the only connector compatible with integrative as well as incremental conceptualization, and capable of introducing clauses with backgrounded as well as foregrounded information. Hence, *weil* spans a range of usage options that includes those of both *denn* and *da*. This suggests the prediction that *weil* is more frequent than *denn* and *da*. Table 2 verifies this for spoken language, implying that *weil* is a jack-of-all-trades: the most widely usable and most easily accessible lexical choice as causal connector.

Finally, consider Figure 7 for an overview of what we discussed in the present section. The left column assumes that conceptualization mode (incremental vs. integrative) correlates with closure point (pre- vs. post-explanans) and level of syntactic integration. The middle column shows the possible combinations of fore-/backgrounding of explanandum or explanans. Underlying this column are two—presumably uncontroversial—combinatorial assumptions: Incremental conceptualization always foregrounds the (leading) explanandum; and in case of integrative conceptualization, the explanans is always expressed by a subordinate clause irrespective of whether it is fore- or backgrounded. The right column summarizes the lexico-syntactic repercussions: eligible connectors and word orders in explanans clauses (to be detailed in Section 4.2).

Figure 7 also reveals that epistemicity and propositionality of the speaker's communicative intention are not allotted a role as (co)determinant of the shape of the explanans clause. We indeed view explanation type as an emergent property reducible to other aspects of the production process.

	Wegener (1999) (1955–1995) N=1095	**Wendekorpus (1993–1996) N=897**	**VERBMOBIL (1993–2000) N=985**
weil	76	81	73
denn	21	14	16
da	3	5	11

Table 2: Relative frequencies of connectors *weil*, *da*, and *denn* in several spoken German corpora. The numbers are percentages of the total number of recorded causal clauses (presumably all "trailing" in the sense of Section 3.1). Sources: Wegener (1999): various corpora, listed in paper. Wendekorpus: Dittmar & Bressem (2005). VERBMOBIL: this study. Between parentheses: period in which recordings were made.

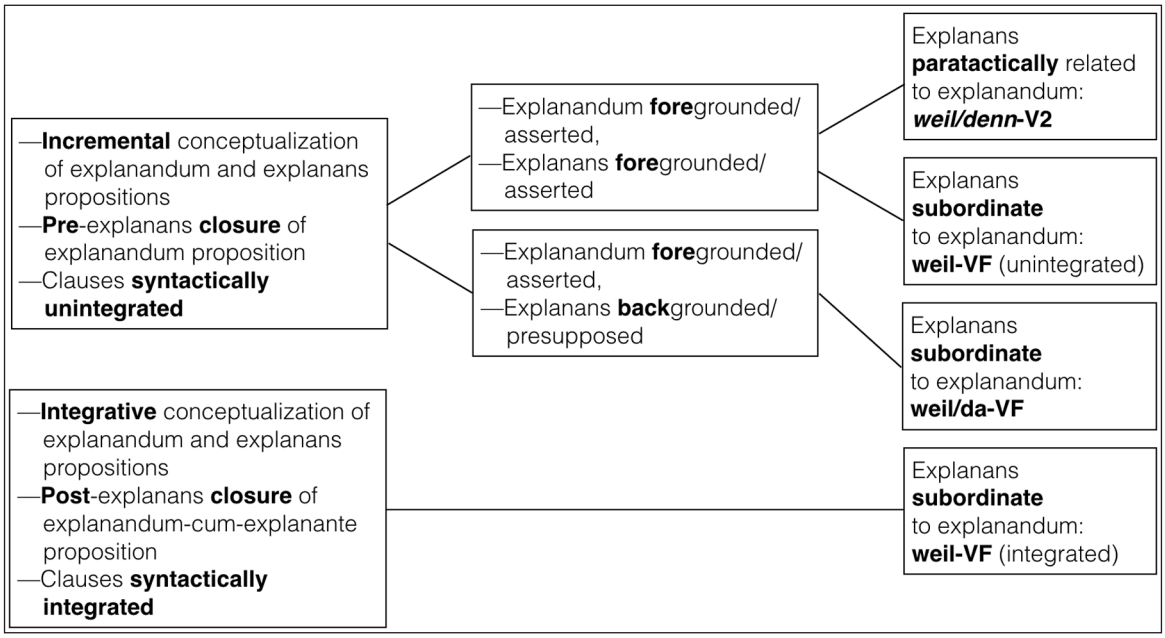

Figure 7: Some important relations between the theoretical notions discussed in Section 4.1.

4.2 Weil-*V2 as a performance phenomenon: Miscoordination between sentence production modules*

The account we now propose for *weil*-V2 may be viewed as, in essence, a combination of the information summarized in Figure 7 and Table 2. The decision tree in Figure 8 highlights the crucial points. The situation denoted at the top in layer A of the figure is that of a speaker who is about to produce either an entire explanandum-cum-explanante sentence (integrative conceptualization) or to add an explanation to an already planned or produced explanandum (incremental conceptualization). The two actions listed in this layer can proceed in parallel (synchronously): activation of the causal connectors in the mental lexicon, and initiation of the conceptualization process that will ultimately engender the proposition(s) to be grammatically encoded. The three connectors compete with one another for insertion in the utterance-under-construction, with *weil* initially being the most likely winner due to its high frequency. This may lead to early insertion of *weil* into the sentence plan. The decision to close the explanandum before or after the explanans (layer D) depends on the planning process: incremental vs. integrative conceptualization and, in the latter case, on the presence of grammatical encoding problems incurred by the explanans (layer C). While the proposition(s) leading to (explanandum and) explanans is/are crystallizing out, their newness or givenness can already affect the syntactic (and prosodic) shape of the utterance, in particular foregrounding (assertion) vs. backgrounding (presupposition) of the explanans: layer E. In parallel, the activation levels of the connectors get modulated: Foregrounding the explanans boosts the activation level of *denn*, backgrounding of the explanans boosts *da*. The (already high) activation level of *weil* is not affected due to its neutrality regarding newness or givenness of the clause it introduces (see below for detail).

What happens when the speaker selects the backgrounding option for the explanans (E2)? This option boosts the activation of *da*-VF which, if able to beat *weil*, will introduce the explanans clause; otherwise *weil* will do (G4 vs. G5). In both cases, the default VF word order will manifest itself.

If no closure has been applied to the explanandum proposition (D2), *weil*-VF is appropriate (G6); if the explanandum was closed, *denn*-V2 is licensed (G2). *Weil*-VF is also licensed

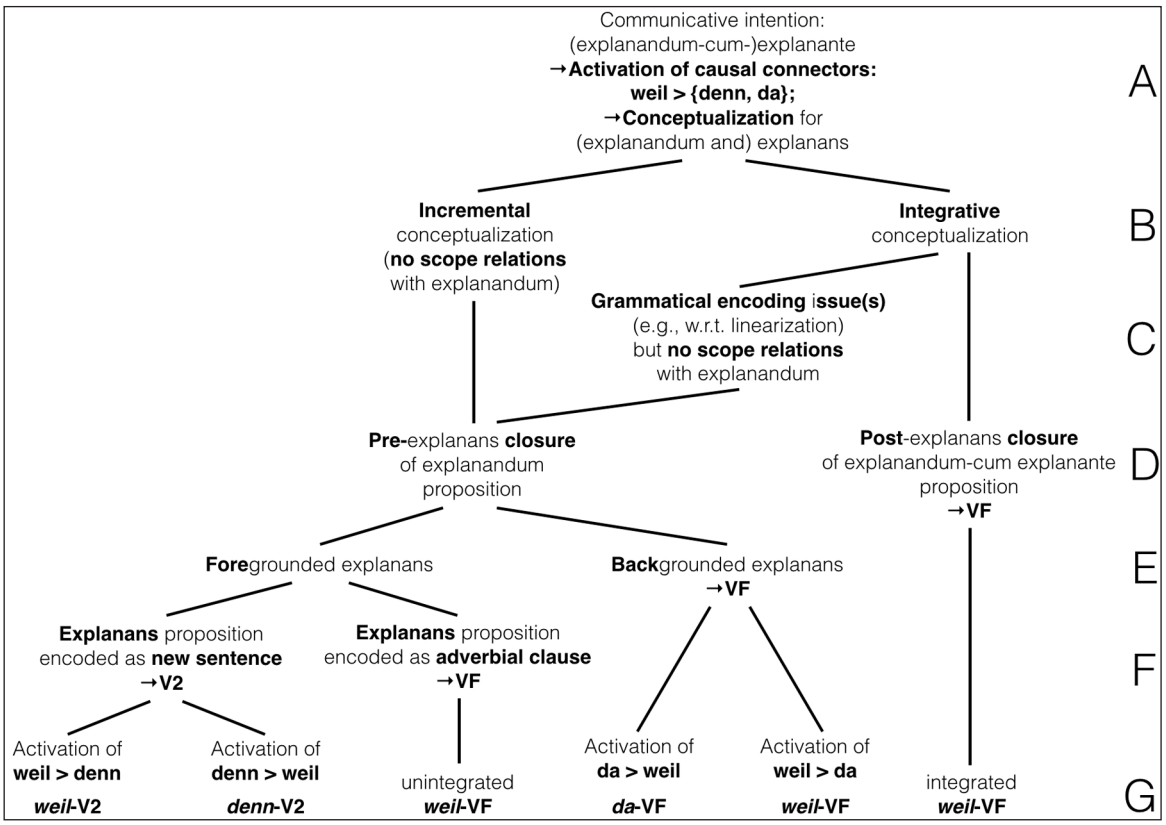

Figure 8: Summary of the proposed speaker-centered account of weil-V2 and related phenomena in the form of a decision tree. Nodes indicate speaker decisions or properties of current conceptual or syntactic structure. Edges below them indicate consequences of decisions or properties. In the text we use the labels A–G to refer to layers of nodes; and nodes within layers are counted from left to right.

in this case (G3), giving rise to an adverbial (subordinate) explanans clause uttered as a kind of afterthought. Reis (2013) treats this option as "unintegrated" *weil*-VF (e.g. in the context of examples such as in (3)).

The miscoordination underlying the ungrammatical *weil*-V2 alternative is depicted in node G1. This choice comes about as a consequence of early—in fact premature—insertion of *weil*, based on its high initial activation before the other connectors have undergone the influence of explanans fore- or backgrounding, and of pre- or post-explanans closure of the explanandum.

This account is not yet complete, though, because it fails to explain why the speaker does not *always* select the "unintegrated *weil*-VF" option (G3). Before addressing this issue, we provide some more details regarding lexical selection of connectors.

Denn and *da* are more specific connectors compared to *weil*. This affects the competition between the three connectors when the grammatical encoder needs one and sends a query to the mental lexicon. In line with much psycholinguistic evidence on the lexicalization process ("word finding"), we assume that the three connectors compete with one another for selection, and that the outcome is determined (partly probabilistically) by the activation levels of the competitors. The current activation depends not only on usage frequency (causing *weil* to be more highly activated than *denn* and *da*), but also on goodness of fit between the query addressed to the mental lexicon and the specific properties associated with candidate lexical items. Lexical items that "match" the query more closely (better fit) receive more activation than lexical items that match the query more poorly. For instance,

if the query lists the property "foregrounding", then extra activation will be gained by *denn* but not by *weil* (which is neutral with respect to prominence level). Specificity thus can raise the activation level of a matching item. In case of a *denn* vs. *weil* contest, if the speaker's communicative intention includes the wish to foreground the explanans proposition, *denn* will get a head start which may enable it to beat the otherwise more frequent, hence more highly activated, connector *weil*. Likewise, the even less frequent *da* has a chance to beat *weil* if the query mentions backgrounding. Finally, also in line with mainstream psycholinguistic assumptions, the winning item sends inhibition (i.e., negative activation) to the competitors, suppressing them to below-threshold activation levels. As a consequence, when *denn* and *da* are defeated, they will be temporarily inaccessible.

The combined effect of frequency-based and specificity-based activation is the following: If *denn* or *da*, in spite of the competition from *weil*, do emerge as winner of the connector competition, they will introduce a clause whose pragmatic properties match the properties associated with them in the mental lexicon. In particular, *denn* will not be forced into combination with a proposition that is to be backgrounded or with a propositional explanation, neither will *da* have to introduce a foregrounded clause. On the other hand, the fact that *weil* is much more frequent than its contestants and easily beats them, readily leads to premature selection—premature in the sense of being uninformed by properties of the upcoming explanans proposition—, and thus running the risk of violating its obligatory VF property. Due to the temporarily subthreshold activation levels of the suppressed *denn* and *da* connectors, these more appropriate competitors cannot present themselves as substitutes. Indeed, we did not find a single V2 explanans where *weil* was replaced by *denn*.

We now turn to the question why speakers often select the ungrammatical V2 option of launching a new sentence (main clause) to accommodate the explanans proposition instead of the well-formed "unintegrated *weil*-VF" alternative. Our answer is based on data in our corpus that signal grammatical encoding problems while planning the explanans (layer C). One of our observations concerns the higher incidence of disfluencies in V2 than in VF explanantia. We attribute these disfluencies to sentence planning problems of any kind—lexical, syntactic, or morphological. Due to such problems, the grammatical encoder may need to mount extra processing capacity and working memory space. One remedy is to close the explanandum proposition and to allocate all available processing capacity to explanans proposition. This "focus (of attention) shift" will lead to rapid decay of the conceptual and syntactic structure built for the explanandum, and to availability of free processing capacity and space for the explanans.[17] This strategy predicts that, other things being equal, explanans clauses that recruit *less* processing capacity are *more* likely to emerge with VF, and vice-versa.

We surmise that this relation underlies the notably higher proportions of non-personal-pronoun subject (PPro–) NPs in *weil*- and *denn*-V2 explanantia compared to their VF counterparts (see Section 3.3, especially Figure 4). Unlike PPro+ subjects, PPro– subjects do not have a privileged placement option such as the Wackernagel position; hence, PPro– subjects have to compete with other clause constituents for a suitable position more frequently than PPro+ subjects. This linearization factor, which increases the processing capacity load to be recruited for PPro– clauses compared to PPro+ explanans clauses, thus may have yielded a higher proportion of explanandum closures. A second factor promoting VF word—or rather, preempting explanandum closures that would give rise to V2—is that personal-pronoun NPs are easy to process referentially (the referents,

[17] Notice that, as already indicated in Section 4.1 and in node C1 of Figure 8, this solution only works if there are no scope dependencies between the two clauses.

internal representations of speaker and addressee, are easily accessible) and lexically (high-frequent, single-word NPs); hence, clauses containing one or more will mobilize less processing capacity than clauses containing full NPs instead. This reduces the need to gain extra working memory capacity by releasing a closed explanandum at the transition to the explanans: Thus, the grammatical encoder can afford to keep the explanandum structure active while processing the relatively simple explanans.

Another condition conducive to efficient management of restrictive cognitive processing capacity arises when the speaker produces a main clause without yet being aware that it will turn into the explanandum of a causative sentence. For example, consider a speaker who, having delivered a verbal description of some event or state of affairs, realizes that this description needs an explanation or justification, and starts planning the explanans clause during or after the delivery of the event description, hence after closure of the explanandum. The speaker's awareness that an explanans is desirable can also come from gestural or other reactions from dialog partners (Eisenberg 2006: 4–5; Diessel & Hetterle 2011: 210). In such cases, the explanandum and explanans clauses are planned incrementally rather than in overlapping time intervals. Consequently, the explanandum is planned, encoded, and closed as a separate sentence; and when the explanans proposition becomes available, it is often too late to interact with the explanandum proposition, with *denn/weil*-V2 as only viable option. Table 1 indeed shows that when prosodic integration decreases (going from top to bottom through the table), the percentage of V2 clauses increases substantially: 44% in the top, 62% in the middle, and 86% in the lower panel.[18] We attribute this rise to increasing likelihood for the explanandum to be closed and cleared from working memory due to increasing asynchrony of the two propositions.

In sum, *weil*-V2 can be accounted for as a performance phenomenon induced by miscoordination between strategies seeking optimal utilization of the processing capacity of the conceptualization, lexicalization, and grammatical encoding mechanisms.

However, the story does not end here.

4.3 An indirect consequence of failing coordination between sentence production modules

The *weil*-V2 engendering miscoordination scenario may have an indirect effect: If it unwinds frequently enough in speakers, it may induce a lexical acquisition/learning process in hearers. Since hearers are speakers themselves, their mental lexicon may develop an association between *weil* and the cognitive V2 scenario consisting of explanandum closure, focus shift toward the explanans, and launching a new sentence. Depending on the frequency of *weil*-V2 perception in the language community, this association can lead to increased *weil*-V2 production by speakers in spite of VF remaining the licensed word order. That is, their *weil*-V2 usage can get triggered by the association instead of by cognitive processing factors of the type to which the V2 engendering scenario presumably owes its existence.

This means we should reckon with the possibility that a non-negligible proportion of the *weil*-V2 clauses in our corpus is due to association-based facilitation of the construction rather than to cognitive load. Indirect support for this suggestion comes from the relatively low correlation we observed between syntactic and prosodic integration (see data summaries at the end of Section 3.3). A relatively high level of prosodic integration is demonstrated not only by the syntactically highly integrated *weil*-VFs but also by the

[18] Percentages calculated by adding, for each level of integration, the two V2 totals, dividing this sum by the sum total at the bottom, and multiplying by 100. For instance, the V2 percentage at higher integration level equals 44 = ROUND((215 + 94)/710*100).

syntactically unintegrated *weil*- and *denn*-V2s. The proportions of "higher level of prosodic integration" observed with *weil*- and denn-V2 clauses are substantial (*weil* 64%, *denn* 59%), although clearly smaller than the very large proportions for VF explanantia (*weil* 80%, *da* 89%).[19]

We account for the high level of prosodic integration observed for *da*-VF clauses by postulating an association between *da* and the following production scenario underlying its use: The intention to communicate an explanans proposition as given/presupposed is always preceded by closure of the foregrounded and asserted explanandum proposition. The original motivation for speakers to apply closure in this speech situation may be similar to the one underlying the production of an explanandum-cum-explanante sentence consisting of two foregrounded/asserted propositions: the fact that the conceptualization process delivers these propositions incrementally, i.e., sequentially. However, presumably due to high usage frequency in the language community, the association between *da* and explanandum closure has become strong enough to get elicited even when the two propositions are conceptualized more or less synchronously. (As *da* always introduces a backgrounded proposition, there is no pressure towards V2 because a main clause must carry assertional illocutionary force.)

Notice that the acquisition process proposed in this section differs from another conceivable acquisition process in language learners/users, viz. that *weil*-V2 could be the result of assimilation of *weil* to *denn*. The acquisition process we propose is secondary to—enabled by—the miscoordination stipulated in Section 4.2. This stipulation has the additional advantage of accounting for the direction of the assimilation: from *weil*-VF to *weil*-V2 instead of vice-versa. The assimilation hypothesis cannot rule out *denn* assimilating to *weil*, with *denn*-VF clauses as result.

Whatever the precise nature of the association formation processes in the mental lexicon, one should expect them to affect not only sentence production and comprehension by language users directly but also, indirectly, their grammaticality/acceptability judgments of self- and other-produced utterances.

5 Discussion

5.1 Summary of corpus data and proposed theoretical framework

We have reported our examination of a collection of about 1,000 spoken German sentences that contain a causal clause ("explanans") introduced by one of the connectors *weil*, *da* or *denn*, positioned immediately after the main clause ("explanandum"). We had extracted these sentences from a corpus of spoken dialogues (VERBMOBIL) recorded between 1993 and 2000. As part of the VERBMOBIL project, the dialogs had been annotated with tags marking major prosodic and syntactic boundaries, and various types of disfluencies (speech pauses, hesitations, self-repairs). We calculated how frequently these annotations occurred in combination with verb-second or verb-final word order in the explanans clause, and with ordering patterns of the first few constituents of the explanans. The observed frequency patterns suggest that an important factor leading speakers to select V2 after *weil* is anticipation, or actual experience, of grammatical encoding problems early on in the *weil* clause. We also noticed that, in this respect, *weil*-V2 clauses resemble the *denn*-V2 clauses in the corpus. Our original conjecture was that sentences with a V2 explanans are somehow harder to plan than sentences with a VF explanans. However, a grammatical phenomenon uniquely associated with *denn*—i.e., not with other

[19] Percentages calculated by dividing the totals in the upper panel of Table 1 by the grant totals in the lower panel, and multiplying by 100. For instance, the *weil*-V2 percentage equals 64 = ROUND((215/336)*100).

connectors—suggested that the V2 explanans could facilitate the processing of complex explanandum-cum-explanante sentences. (In the German linguistic literature, this *denn* property is known as parataxis; see citations at the onset of Section 4.1). *Denn* affords the speaker a "divide and conquer" strategy that splits the parallel-interactive planning of a complex sentence into the sequential planning two simpler sentences, first the explanandum, then the explanans sentence.

Could V2 after *weil* be another manifestation of the same *divide et impera* strategy? If so, the V2 clause introduced by *weil* should not be viewed as a clause subordinate to the preceding explanandum but as a second independent sentence featuring V2 as the default ordering pattern for main clauses. This presupposes that the explanandum proposition has undergone closure ("freezing"). In this case, V2 is the only possible word order option within the explanans clause. Otherwise (i.e., in the absence of closure), the explanandum proposition can be extended with an explanans proposition, enabling the grammatical encoder to express the latter proposition as a VF clause attached to the explanandum clause (provided the explanans is foregrounded/asserted). The latter scenario is the more likely, the simpler the planning process for the explanans. This predicts, for instance, a higher incidence of VF in clauses with personal-pronoun subjects occupying the preferred Wackernagel position. However, when grammatical encoding problems with respect to the explanans are expected or experienced, the encoder can deploy the divide and conquer strategy, and cast the explanans in the V2 mold.

Note that closure of the explanandum in case of *denn* is not an encoder operation specifically associated with this connector. If indeed *denn* is paratactically related to the explanandum, with the *denn* clause starting up an independent new sentence, then explanandum closure is just the standard operation the encoder executes when terminating *any* sentence prior to embarking on the next sentence in the discourse. By virtue of the linguistic arguments put forward in Section 4.1, we stipulate that this also holds for *weil* when followed by a V2 clause.

The proposed account of *weil*-V2 proposed in Section 4 assumes that VF is the only word order option licensed by *weil*. The reason it is regularly followed by a V2 clause has to do with *weil*'s high frequency in language use—much higher than that of *denn* and *da*—and with the fact that it does not impose pragmatic restrictions on the explanans clause. The *weil* clause may express a foregrounded/new/asserted proposition or a backgrounded/old/presupposed proposition. This contrasts both with *denn*, which introduces an assertion, and with *da*, which introduces a presupposition. The broad usability of *weil* conspires with its high frequency and high resting level of activation to make it an easily accessible and attractive lexical candidate to express causation or causal inference. However, if selected early, its VF property may clash with V2 word order in an explanans clause that is currently being encoded as an independent new sentence—due to a currently high cognitive processing load (Section 4.2) or to a strong association between *weil* and the closure-with-focus-shift scenario (Section 4.3); and since the early, premature selection of *weil* has led to suppression of the appropriate alternative *denn*, the latter cannot rescue the imminent violation of *weil*'s VF projection constraint.

The larger the number of language users in whose mental lexicon an association develops between *weil* and the V2 engendering cognitive scenario, the lower the threshold for *weil*-V2 to get activated and to affect speech situations without, or with only mild, cognitive processing load. A trend in this direction within the language community is expected to render the *weil*-V2 construction acceptable in spoken language use.

In the remainder of this section, we compare our account with extant theoretical frameworks, and confront it with additional empirical findings.

5.2 Some linguistic and psycholinguistic repercussions

Separate weil-*V2 and* weil-*VF lexical entries?* In the final section of her paper, Reis (2013) expresses dissatisfaction with an important theoretical implication of her analysis of *weil*-V2 phenomena: the requirement that the lexicon of modern German contains *two* lexical items for the conjunction *weil*, which project a VF clause and a V2 clause, respectively. In view of the fact that the meanings associated with the two *weil* items are identical, she would prefer "to posit only one lexical entry and to derive the distributional differences (syntactic as well as interpretive) from the regular impact of other factors".[20]

The proposed model preempts the necessity of postulating two *weil* entries in the mental lexicon of speakers of German who regularly use VF and V2.[21] According to the model, V2 constituent order in the clause following *weil* is not licensed by *weil* but by word order rules for main clauses (see also Footnote 14). According to the model, *weil*-V2 emerges precisely because the explanans proposition has "escaped" from the *weil*-clause. An additional advantage is that no selection process between two *weil* entries is required.

Explanandum closure also in the hearer? There is mounting psycholinguistic and neurobiological evidence of close collaboration, if not substantial overlap, between the cognitive processing modules responsible for the grammatical aspects of sentence production and comprehension, that is, between the grammatical encoder and the grammatical decoder (Kempen 2000, 2014; Segaert et al. 2012). Hence, the comprehension system of a hearer may "know" that the production of a V2 explanans clause by the interlocutor must have been preceded by closure of the proposition expressed in the explanandum. In response, the hearer will also close his/her interpretation of the explanandum, thereby restricting the interpretation of the ensuing explanans clause to one resulting from incremental conceptualization, i.e., an epistemic one (FCI or BCI).

This reasoning predicts the results of an interesting experiment reported by Antomo & Steinbach (2010: 25ff). In a test with printed materials, they presented their participants (a group of German university students) with sentences such as (8a) and (8b). These sentences are ambiguous in that they leave open which of the two described events is the physical cause of the other one. They found that *weil*-VF clauses such as in (8a) elicited a large majority (72.3%) of propositional explanations (airbag inflation as cause), whereas *weil*-V2 clauses as in (8b) elicited only a small majority (55.5%) of epistemic explanations (airbag inflation as effect).

(8) a. Es hat einen Unfall gegeben, weil der Airbag aufgegangen
there has an accident been because the airbag inflated
ist. (VF)
has
'There has been an accident because the airbag inflated.'

b. Es hat einen Unfall gegeben, weil der Airbag ist aufgegangen. (V2; Antomo & Steinbach 2010: 32)

The authors put forward a detailed pragmatic-linguistic account of why *weil*-V2, in contrast with *weil*-VF, allows a reversal of the understood direction of causation. This account (partly going back to Blakemore & Carston 2004) postulates a comprehension system based on specific "pragmatic processing units" and "processing stages" (for a brief summary,

[20] "[Die Wunsch-Analyse dafür wäre unbedingt, dann auch] nur einen Lexikoneintrag anzusetzen, und die Distributionsunterschiede (syntaktische wie interpretative) aus dem regulären Einwirken anderer Faktoren herzuleiten." (Reis 2013: 255)

[21] Frequency counts of the conjunction *denn* have shown great variation with respect to the incidence of this conjunction in German dialects. In Southern Germany, in particular, frequencies were often low. *Weil*-VF is the most widely used causal conjunction; *weil*-V2 is on the rise. Cf. Pasch (1997) and Wegener (2000).

see Section 2). We make the—presumably uncontroversial—assumption that the grammatical decoder parses the VF clause in (8a) as resulting from integrative conceptualization. This engenders the propositional explanation where airbag inflation is cause of the accident—the students' preferred reading (72.3%). The same standard parsing strategy fails when the decoder detects V2 word order in the explanans. However, the grammatical decoder recognizes in the *weil*-V2 structure a pattern that the grammatical encoder generates when the communicative intention necessitates incremental conceptualization, namely, explanandum closure. Therefore, the decoder also closes the explanandum proposition and takes the relation between explanandum and explanans as one of justification (FCI or BCI) rather than CC (a causal chain; see Section 3.2). This implies an epistemic rather than a propositional explanation, with the inflated airbag as evidence/argument for the inference about the accident. However, given knowledge of the modern world, the hearer cannot decide between airbag inflation as cause or as effect of the accident (FCI versus BCI). Hence, the participants could not do much better than making a random choice (55.5%).

Verb-second in adverbial clauses with other connectors? As is well-known, V2 word order can also be projected by certain non-causal subordinating conjunctions. Examples are *obwohl* 'although' (concessive) and *während* 'whereas' (adversative), often preceded and/or followed by a short prosodic break. However, other subordinating conjunctions (temporal, conditional, instrumental, resultative) do not license V2. (To illustrate in terms of VERBMOBIL: Of the 71 *obwohl* tokens in the corpus, 47 preceded a V2 clause (66%), and 34 of the latter were tagged with a disfluency mark after the connector.)

It does not seem farfetched to generalize the theoretical framework we developed for causal V2 clauses to concessive and adversative clauses: Trailing *obwohl* and *während* clauses with V2 assert independent propositions, on a par with *weil/denn*-V2 clauses; and the main clauses preceding them are also independent speech acts. If the generalization holds, we also need to assume that concessive and adversative V2 clauses are accompanied by closure of the main clause proposition. The generalization also forces us to postulate that the other group of connectors (temporal, conditional, etc.) do *not* close their main clause. Stated differently, they keep the main clause proposition open, susceptible to modification by additional conceptual content (integrative conceptualization); and the result is *one* asserted proposition spanning the entire sentence (similar to the single asserted proposition conceived during planning an integrated causative sentence where the causal clause modifies the meaning of the main clause). We are confident that the first part of the generalization—obligatory closure of the proposition preceding *obwohl* and *während*—is correct, viewing the disfluencies that tend to accompany them.

Whether the second part holds (i.e., no closure before temporal, conditional, etc., connectors), is an open empirical issue. An indirect argument in its favor is derivable from the VERBMOBIL data concerning the connector *sodass* 'so that'—the "resultative" VF connector which expresses a relation similar to causality, although in a direction opposite to that of *weil*-VF. (In the corpus, the connector is spelled as two contiguous words: *so daß*—according to old German spelling.) The corpus contains 63 exemplars of *sodass*, all unambiguously classifiable as VF. Remarkable, and in strong contrast with *obwohl*, is their high level of prosodic integration: Only 6 *sodass*-VF clauses were realized with a disfluency around the connector. Moreover, 46 out of the 63 *sodass* tokens (73%) had a personal pronoun subject in Wackernagel position (PPro + W +; the remaining 27% were all PPro–). These proportions strongly resemble those of *weil/da*-VF depicted in Figure 4. On the assumption that *sodass* clauses represent foregrounded/new information within the complex sentence, it follows that the underlying production process is similar to that depicted in the rightmost branch of Figure 8 (from layer A to node G6).

Although this scenario indeed does not include closure of the main clause preceding *sodass*, it leaves open the possibility for the *sodass* clause to be within the scope of operators in the main clause. However, we have not been able to find—or construct—any well-formed example of such a structure, suggesting that the apparent prohibition on cross-clause scope relations with *sodass* has another origin than pre-*sodass* closure of the main clause. Here are two options (both presuming that *sodass* denotes a causal relationship). The first one proceeds from the assumption that speakers who introduce scope relations between clauses, deploy logical reasoning as part of the conceptualization process, hence manipulate propositions. This would rule out using *sodass* if its lexical entry would specify that it does not connect propositions but events referred to by propositions (i.e., defining *sodass* as a referential rather than inferential connector; cf. the familiar distinction between cause/effect vs. argument/conclusion). However, even if true, this would merely describe the observation, not explain it. The second speculation: *Sodass* clauses are always foregrounded/asserted whereas the preceding main clauses are virtually always backgrounded/presupposed. This would imply integrative conceptualization and preclude the divide-and-conquer solution of splitting the complex sentence into a pair of independent sentences. Consequently, the frequency of this solution would be too low to trigger the learning process sketched in Section 4.3: associating *sodass* with main clause closure. However, this hypothetical account is incomplete without a reason why *sodass*, in contrast with the almost equally frequent *obwohl*, is rarely preceded by a foregrounded/asserted main clause. Hence, we have to leave unanswered the question why there is no *sodass*-V2 phenomenon.

Finally, we wish to note that the theoretical framework developed in this paper should be viewed as a set of hypotheses suggested by the corpus data, and needs additional empirical support from experiments that test predictions derived from it. This requires experiments manipulating the speaker's communicative intention and the cognitive load imposed by (un)problematic explanans clauses as independent variables, and assess the tendency for speakers to use verb-second or verb-final word order as dependent variable, preferably in conjunction with other behavioral and neurobiological measurements. Empirical work of this type is likely to have impact beyond the domain of a handful of lexical items, on the study of language production generally.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgement

We are greatly indebted to four reviewers for detailed and constructive comments.

Appendices

A.1 Some weil*-V2 sentences from the VERBMOBIL corpus*

The annotation symbols are explained in Appendix A.2 below.

Weil-V2, with Left Dislocation

(9) <#Klopfen> ja, +/wei=/+ weil November, da kann 's schon
yes because November there kann it already
<:<#> geschneit:> haben <!1 hamm>. <A> w"urde <!1w"udd> sagen
snowed have would say
<!1 sang>, wir <:<#> fliegen:>.
we fly
'Yes, because in November it may already have snowed. [I] would say we fly.'

Major intonational break ("?"), hesitation within *weil*-V2

(10) und <"ahm> wie ist <!1 's'> es ab dem #siebenundzwanzigsten?
and how is it from the seventheenth
w<Z>eil <Ger"ausch> bis zum <A> August, da h"atte ich
because until August there would.have I
auch noch viel Zeit.
also still much time
'And how about from the seventeenth onward? Because until August I still have ample time, too.'

Major intonational break (".") , with hesitations before and after *weil*-V2

(11) ja<Z>, dann<Z> <#> <P> ist das doch an sich schon<Z> perfekt
yes then is that indeed in itself already perfect
dieses Treffen, dann k"onnen wir doch<Z> <Schmatzen> uns<Z>
this meeting then could we indeed us
<"ah> in dieser Woche, also, in diesen #f"unf Tagen<Z> <A> dort
in this week that.is in these five days there
in ~Frankfurt treffen<Z>. <Schmatzen> <"ah> weil<Z> +/ im<Z>%/+
in Frankfurt meet because in.the
<P> also, fr"uher geht bei Ihnen auch nicht im<Z> Juli? sonst
well earlier goes with you also not in.the July otherwise
k"onnt' ich da eventuell gesch"aftlich noch was verschieben,
could I there perhaps business-related still something move.up
<A> <"ah> falls es Ihnen dann besser pa"st. <#Klopfen>
if it you then better suit
'Yes, then this meeting is already perfect in itself, then we can meet this week, that is, in these five days, in Frankfurt. Because, well, earlier in July is impossible for you? Otherwise, I could move up some other business appointment if that suits you better then.'

Smooth *weil*-V2 with conditional clause in first position

(12) ja, dann<Z> w"ar' vielleicht <!1 v'lleicht> am besten das <!1 des>
yes then would perhaps at.the best the
Hotel, was zentral ist <!1 is'>, weil wenn wir abends tanzen
hotel that central is because if we in.the.evening dance
gehen wollen, <A> w"ar' das <!1 des> vielleicht <!1 v'lleicht> am
go want would that perhaps at.the
besten. <Schmatzen> <A> und <!1 un'> mit<Z> Hallenbad ist <!1 's'>
best and with indoor.pool is
vielleicht <!1 v'lleicht> auch nicht <!1 nich'> schlecht. <A> dann
perhaps also not bad so
werde <!1 werd'> ich mich um die<Z> Buchung von dem Hotel
will I myself about the booking of the hotel
k"ummern.
care
'Yes, then the hotel that is located in the center would perhaps suit us best, because if we want to go out for dancing in the evening, this would perhaps suit us best. And with the indoor pool is also not bad. So, I will take care of the hotel booking.'

Weil-V2 with conditional clause; hesitations before and after *weil*

(13) <#Klopfen> <Schmatzen> <A> genau<Z>. <P> das wirft nat"urlich
indeed that brings of.course
bei der Planung einige Probleme<Z> auf<Z>, weil <"ahm> <P>
for the planning some problems up because
wenn wir das block. <P> seminarm"a"sig abhalten w"urden, ginge
if we that block seminar.like conduct would would.go
das nat"urlich etwas schneller<Z>. <Schmatzen> haben Sie denn
that of.course a.bit faster have you so
Termine<Z>, wo Sie absolut nicht k"onnen? <#Klopfen>
time.slots where you absolutely not can
'Indeed, that causes some planning problems, because if we conduct it as a block seminar, it would be a bit faster. Do you have time slots where you cannot make it at all?'

Smooth *weil*-V2, with personal-pronoun subject in Wackernagel position; reply to a question by the dialogue partner

(14) <#Klopfen> <A> <"ah> da klappt es nicht <!1 nich'>. <"ahm>
then suits it not
<Ger"ausch> geht es am Nachmittag? <#Klopfen> <Schmatzen>
goes it in.the afternoon
schlecht, weil ich <!2 wei'ch> hab' um #zwei Kolloquium.
bad because I have at two colloquium
<#Rascheln> <A> <Schmatzen> lass' uns mal <!1 ma'> lieber nach
let us just better for
'nem andern Tag kucken.
an other day look
'That doesn't suit me, because at two I have a colloquium. Let's better look at another day.'

Two smooth *weil*-V2 cases, once with late personal-pronoun subject, once without any personal-pronoun subject

(15) ah, ja, danke, dann hab' ich mich wohl verh"ort. <A>
oh yes thanks thus have I myself probably misheard
Moment,<Schmatzen> <A> na, das ist <!2 des is'> bei mir
one.moment well that is with me
aber <!1 awa> sehr schlecht, weil in der Woche hab' <!1 hob>
however very bad because in this week have
ich <!1 i'> schon einen Termin in ~Frankfurt. <A>
I already an appointment in Frankfurt
<Schmatzen> wie <:<#Rascheln> w"ar':> <:<#Rascheln> 's:> denn
how would it then
bei Ihnen dann im <:<#Rascheln> Dezember:>, weil
with you then in.the December because
so ist <!1 is'> bei mir schon <:<#Rascheln> alles:>
in.this.manner is with me already everything
<:<#Rascheln>ziemlich:> voll?
rather full

'Oh yes, thanks, then I probably have misunderstood. One moment, well, however, that suits me very badly, because in this week, I already have an appointment in Frankfurt. How would it be then for you in December, because in this manner everything is rather booked in my agenda.'

A.2 Most important symbols used in the transliterations (quoted and translated from the German VERBMOBIL documentation)

Word reductions and pronunciation variants:
haben <!1 hab´>, und dann <!2 un´a´>
have and then

Names and places: Herr ~Huber, in ~Frankfurt
Mister Huber in Frankfurt

Abbreviations: $U-$S-$A, $U-Bahn
U.S.A. u-train

Foreign words: <*ENG>strange, <*ITA>bella

Numbers: #zwei-und-zwanzigsten
two and twenty 'twenty-two'

Self-corrections, repetitions, false starts: -/im Ja=/– also, +/Treb=/+ Traube . . .
in.the so grape

Hesitation types:

<"ah>	purely vocalic articulation irrespective of vowel quality,
<"ahm>	vocalic articulation + nasal articulation,
<hm>	purely nasal articulation,
<h"as>	articulation not belonging to any of the above-mentioned classes,
<Z>	hesitation (may also occur within a word if the vowel is stretchted, e.g. gera<Z>de), and
<P>	pause (a longer hesitation than <Z>; it can only occur between words).

The following tags were ignored for the calculations reported in the paper because arguably they are the consequence of physiological rather than cognitive processes:

<A>	breathing,
<Schmatzen>	lip smack,
<Räuspern>	clearing the throat,
<Schlucken>	swallowing,
<Husten>	coughing,
<Lachen>	laughing, and
<Ger"ausch>	any other sounds, including sounds caused by the technical equipment.

Punctuation symbols (comma, full stop, question mark):
The symbols full stop, question mark and comma are available to mark normal sentence parts. However, punctuation often causes problems in spontaneous speech because "correct" sentences are often missing.

Full stop: In dubious cases, reasons for inserting a full stop may be: grammatical structure, intonation, pause, breathing, beginning of a new thought.

Example:

... so, guten Tag. <A> <"ah> mein Name ist ~ J"ansch. gestern hatte ich schon ...
so good day <breathing> my name is Jänsch yesterday had I already

Question mark: In dubious cases, reasons for inserting a question mark may be: Wh-words, sentence structure, intonation, context.

Comma: In dubious cases, reasons for a comma may be: grammar, a particle introducing a subordinate clause, intonation, connected thoughts, phrases separated by breathing or pause.

(16) a. ich könnte schon, mittwochs hab´ ich immer Zeit.
I could even Wednesday have I always time
'I suits me, on Wednesday I always have time.'

b. ich könnte schon. <A> <P> mittwochs hab´ ich immer Zeit.
I could even Wednesday have I always time
'I suits me, on Wednesday I always have time.'

c. <"ahm> morgen, Freitag, <h"as> wie ich seh', <"ah> mu"s
tomorrow Friday as I see have.to
ich feststellen, da"s ich <"ah> "uberhaupt keine Zeit hab' ....
I realize that I at.all no time have
'Tomorrow, Friday, as I just see, I have to admit that I have no time at all.'

Remark: Punctuation symbols can only be preceded by lexical items or comments. Any other event, including breathing, hesitations or completely ununderstandable sounds should be placed after the punctuation symbol. Commas cannot be inserted at the end of a turn.

References

Antomo, Mailin. 2012. Interpreting embedded verb second: Causal modifiers in German. In Bert Le Bruyn & Kathrin Linke (eds.), *Proceedings of ConSOLE XVII*, 27–51. Nova Gorica, Slovenia (2009). Available at http://media.leidenuniv.nl/legacy/console17-antomo.pdf.

Antomo, Mailin & Markus Steinbach. 2010. Desintegration und Interpretation. *Weil–V2-Sätze an der Schnittstelle zwischen Syntax, Semantik und Pragmatik. Zeitschrift für Sprachwissenschaft* 29. 1–38. DOI: http://dx.doi.org/10.1515/zfsw.2010.001

Bisiada, Mario. 2013. Changing conventions in German causal clause complexes: A diachronic corpus study of translated and non-translated business articles. *Languages in Contrast* 13. 1–27. DOI: http://dx.doi.org/10.1075/lic.13.1.01bis

Blakemore, Diane & Robyn Carston. 2004. The pragmatics of sentential coordination with *and. Lingua* 115. 569–589. DOI: http://dx.doi.org/10.1016/j.lingua.2003.09.016

Bock, Kathryn, David E. Irwin, Douglas J. Davidson & Willem J.M. Levelt. 2003. Minding the clock. *Journal of Memory and Language* 48. 653–685. DOI: http://dx.doi.org/10.1016/S0749-596X(03)00007-X

Breindl, Eva. 2009. Fehler mit System und Fehler im System. Topologische Varianten bei Konnektoren. In Marek Konopka & Bruno Strecker (eds.), *Deutsche Grammatik – Regeln, Normen, Sprachgebrauch*, 274–306. Berlin: de Gruyter.

Burger, Susanne, Karl Weilhammer, Florian Schiel & Hans G. Tillmann. 2000. Verb-mobil data collection and annotation. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation*, 537–549. Berlin: Springer.

Diessel, Holger & Katja Hetterle. 2011. Causal clauses: A cross-linguistic investigation of their structure, meaning and use. In Peter Siemund (ed.), *Linguistic universals and language variation*, 21–52. Berlin: de Gruyter.

Dittmar, Norbert & Jana Bressem. 2005. Syntax, Semantik und Pragmatik des kausalen Konnektors *weil* im Berliner 'Wendekorpus' der neunziger Jahre. In Johannes Schwitalla & Werner Wegstein (eds.), *Korpuslinguistik deutsch: synchron – diachron – kontrastiv*, 99–123. Tübingen: Niemeyer.

Drach, Erich. 1937. *Grundgedanken der deutschen Satzlehre*. Frankfurt a.M.: Diesterweg. Reprinted 1963.

Dryer, Matthew S. 1996. Focus, pragmatic presupposition and activated propositions. *Journal of Pragmatics* 26. 475–523. DOI: http://dx.doi.org/10.1016/0378-2166(95)00059-3

Eisenberg, Peter. 2006. Grundriss der deutschen Grammatik, Band 2: Der Satz. Stuttgart/Weimar: Metzler.

Freywald, Ulrike. 2009. Kontexte für nicht-kanonische Verbzweitstellung: V2 nach *dass* und Verwandtes. In Veronika Ehrich, Christian Fortmann, Ingo Reich & Marga Reis (eds.), *Koordination und Subordination im Deutschen*, 113–134. Hamburg: Buske (= Linguistische Berichte, Sonderheft 16).

Freywald, Ulrike. 2010. *Obwohl* vielleicht war es ganz anders: Vorüberlegungen zum Alter der Verbzweitstellung nach subordinierenden Konjunktionen. In Arne Ziegler (ed.), *Historische Textgrammatik und Historische Syntax des Deutschen*: *Traditionen, Innovationen, Perspektiven*, 55–84. Berlin: de Gruyter.

Gaumann, Ulrike. 1983. *"Weil die machen jetzt bald zu". Angabe- und Junktivsatz in der deutschen Gegenwartssprache.* Göppingen: Kümmerle.

Günthner, Susanne. 1993. *"… weil man kann es wissenschaftlich untersuchen"* — Diskurspragmatische Aspekte der Wortstellung in WEIL-Sätzen. *Linguistische Berichte* 143. 37–59.

Hess, Wolfgang J., Klaus J. Kohler & Hans G. Tillman. 1995. The Phondat-Verbmobil speech corpus. In *Proceedings of Eurospeech'* 95. 863–866. Madrid, Spain.

Hobbs, Jerry R., Mark E. Stickel, Douglas E. Appelt & Paul Martin. 1993. Interpretation as abduction. *Artificial Intelligence* 63. 69–142. DOI: http://dx.doi.org/10.1016/0004-3702(93)90015-4

Höhle, Tilman N. 1986. Der Begriff 'Mittelfeld': Anmerkungen über die Theorie der topologischen Felder. In Walter Weiss, Herbert E. Wiegand & Marga Reis (eds.), *Akten des VII. Internationalen Germanistenkongresses*, 329–340. Göttingen 1985, Band 3. Tübingen: Niemeyer.

Keller, Rudi. 1995. The epistemic *weil*. In Dieter Stein & Susan M. Wright (eds.), *Subjectivity and subjectivisation: Linguistic perspectives*, 16–30. Cambridge UK: Cambridge University Press. DOI: http://dx.doi.org/10.1017/cbo9780511554469.002

Kempen, Gerard. 2000. Could grammatical encoding and grammatical decoding be subserved by the same processing module? *Behavioral and Brain Sciences* 23. 38–39. DOI: http://dx.doi.org/10.1017/S0140525X00402396

Kempen, Gerard. 2014. Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding. *Neuroinformatics* 12. 111–142. DOI: http://dx.doi.org/10.1007/s12021-013-9191-4

Kempen, Gerard & Karin Harbusch. 2012. A frequency-based argument for SOV as unmarked word order in German and Dutch. In *Proceedings of 34. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft*, 189. Frankfurt a.M., Germany. [Abstract].

Kempen, Gerard & Karin Harbusch. 2015. In spoken German and Dutch, high-frequent finite verbs populate main clauses more densely than subordinate clauses, but much less so in spoken English: A corpus-linguistic study into VO versus OV word order.

In *Proceedings of 37. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft*, 123–124. Leipzig, Germany. [Abstract].

Lee, Eun-Kyung, Sara Brown-Schmidt & Duane G. Watson. 2013. Ways of looking ahead: Hierarchical planning in language production. *Cognition* 129. 544–562. DOI: http://dx.doi.org/10.1016/j.cognition.2013.08.007

Pasch, Renate. 1983. Untersuchungen zu den Gebrauchsbedingungen der deutschen Kausalkonjunktionen *da, denn* und *weil. Linguistische Studien des Zentralinstituts für Sprachwissenschaft (Berlin), Reihe A: Arbeitsberichte* 104.

Pasch, Renate. 1997. *Weil* mit Hauptsatz-Kukucksei im *denn*-Nest. *Deutsche Sprache* 25. 252–271.

Pasch, Renate, Ursula Braußе, Eva Breindl & Ulrich H. Waßner. 2003. *Handbuch der deutschen Konnektoren. Linguistische Grundlagen der Beschreibung und syntaktische Merkmale der deutschen Satzverknüpfer (Konjunktionen, Satzadverbien und Partikeln)*. Berlin: de Gruyter.

Reis, Marga. 2013. "*Weil*-V2"-Sätze und (k)ein Ende? Anmerkungen zur Analyse von Antomo & Steinbach (2010). *Zeitschrift für Sprachwissenschaft* 32. 221–262. DOI: http://dx.doi.org/10.1515/zfs-2013-0008

Scheutz, Hannes. 2001. On causal clause combining. In Margret Selting & Elizabeth Couper-Kuhlen (eds.), *Studies in Interactional Linguistics*, 111–139. Amsterdam: Benjamins. DOI: http://dx.doi.org/10.1075/sidag.10.07sch

Schönenberger, Manuela. 2000. The acquisition of verb placement in Lucernese Swiss German. In Marc-Ariel Friedemann & Luigi Rizzi (eds.), *The acquisition of syntax: Studies in comparative developmental linguistics*, 293–319. Harlow (Essex): Longman.

Schönenberger, Manuela. 2010. The surprisingly high proportion of Verb Second in causal adverbial clauses with *weil* 'because' in child Swiss German. *Lingua* 120. 2589–2604. DOI: http://dx.doi.org/10.1016/j.lingua.2010.06.009

Segaert, Katrien, Laura Menenti, Kirsten Weber, Karl M. Petersson & Peter Hagoort. 2012. Shared syntax in language production and language comprehension – An fMRi study. *Cerebral Cortex* 22. 1662–1670. DOI: http://dx.doi.org/10.1093/cercor/bhr249

Stede, Manfred. 2008. Connective-based local coherence analysis: A lexicon for recognizing causal relationships. In *Semantics in Text Processing – STEP 2008 Conference Proceedings*, 221–237. London: College Publications. DOI: http://dx.doi.org/10.3115/1626481.1626499

van de Velde, Maartje, Gerard Kempen & Karin Harbusch. 2015. Dative alternation and planning scope in spoken language: A corpus study on effects of verb bias in VO and OV clauses of Dutch. *Lingua* 165. 92–108. DOI: http://dx.doi.org/10.1016/j.lingua.2015.07.006

Wahlster, Wolfgang (ed.). 2000. *Verbmobil: Foundations of speech-to-speech translation*. Berlin: Springer. DOI: http://dx.doi.org/10.1007/978-3-662-04230-4

Wegener, Heide. 1993. *Weil: das hat schon seinen Grund* — Zur Verbstellung in Kausalsätzen mit *weil* im gegenwärtigen Deutsch. *Deutsche Sprache* 21. 289–305.

Wegener, Heide. 1999. Syntaxwandel und Degrammatikalisierung im heutigen Deutsch. Noch einmal zu *weil*-Verbzweit. *Deutsche Sprache* 27. 3–26.

Wegener, Heide. 2000. *Da, denn* und *weil*—der Kampf der Konjunktionen: Zur Grammatikalisierung im kausalen Bereich. In Rolf Thieroff, Matthias Tamrat, Nanna Fuhrhop & Oliver Teuber (eds.), *Deutsche Grammatik in Theorie und Praxis*, 92–108. Tübingen: Niemeyer.

How to cite this article: Kempen, Gerard and Karin Harbusch. 2016. Verb-second word order after German *weil* 'because': Psycholinguistic theory from corpus-linguistic data. *Glossa: a journal of general linguistics* 1(1): 3. 1–32, DOI: http://dx.doi.org/10.5334/gjgl.46

Published: 01 April 2016

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS