RESEARCH

What usage can tell us about grammar: Embedded verb second in Scandinavian

Spencer Caplan and Kajsa Djärv
University of Pennsylvania, Philadelphia, PA, US
Corresponding author: Spencer Caplan (spcaplan@sas.upenn.edu)

This paper uses large-scale data extracted from a series of Swedish corpora to investigate the factors responsible for conditioning the choice of (optional) embedded V2 in Swedish. Embedded V2 has been argued to represent a more general kind of syntactic optionality found across languages: syntactic structures typically found in matrix clauses, but which are also available in certain types of embedded environments (so called Main Clause Phenomena). While the received view, going back to Hooper & Thompson (1973), is that the availability of main clause syntax has a semantic-pragmatic correlate in the presence of Illocutionary Force, pinpointing exactly what this amounts to has remained an open problem. Through statistical analysis of the Swedish corpus data, combined with results from a semantic-inference task, we are able to falsify certain previous (theoretical and empirical) claims about the distribution and interpretation of embedded V2. We additionally evaluate, and find no evidence to support, a processing or usage-based view of optionality in embedded V2. We argue instead that the interpretive notion driving the distribution of embedded V2 is discourse novelty; whether the embedded proposition is treated as discourse-old or new information. We argue that embedded V2 is licensed in contexts where p is discourse novel. While this is fundamentally a pragmatic notion, it is nevertheless tightly constrained by both lexical-semantic properties of the matrix predicate and other aspects of the grammatical context. An important methodological consequence of this work is that by looking at particular interactions of lexical and grammatical contexts, statistical analysis of usage data can be used to test specific predictions made by syntactic and semantic theory.

1 Background

1.1 Introduction

In this paper, we investigate a type of syntactic optionality found across languages: syntactic structures typically found in matrix clauses, but which are also available—although apparently not obligatory—in certain types of embedded environments. Since Hooper & Thompson (1973), building on seminal work by Emonds (1970), the received view is that such Main Clause Phenomena [MCP] are licensed by the kinds of interpretive properties typically associated with matrix clauses. In particular, the received view is that MCP are available in contexts associated with Illocutionary Force. However, pinpointing exactly what this association amounts to has proven a serious challenge for both theoretical and experimental work on this topic (see among many others Hooper 1975; Green 1976; Wechsler 1991; Holmberg & Platzack 1995; Gärtner 2000; 2002; Truckenbrodt 2006; Julien 2009; Wiklund et al. 2009; Bentzen 2010; Gärtner & Michaelis 2010; Haegeman & Ürögdi 2010; Wiklund 2010; Aelbrecht et al. 2012; Haegeman 2012;

Jensen & Christensen 2013; Haegeman 2014; Julien 2015; Woods 2016a; b; Miyagawa 2017; Djärv et al. 2017). This paper presents new quantitative data addressing this question in the context of embedded V-to-C movement, or Verb Second—a type of MCP found across a variety of languages, including Mainland Scandinavian and several other Germanic languages. From this data, we argue that embedded V2 [EV2] is only licensed in contexts where the embedded proposition p is introduced into the conversation as entirely new information. This is to say that EV2 is unavailable in contexts where p has been previously discussed by the speaker and the hearer, regardless of whether or not p is mutually agreed on.

Scandinavian EV2 raises a number of questions for the study of syntactic optionality. While previous work on EV2 has reported judgments pointing to potential semantic-pragmatic factors driving the choice of EV2 vs. V-in situ, consensus has yet to be reached as to what precisely those factors are. One possibility, which we consider in this paper, is that any interpretive correlates are only apparent, and that the choice is in fact primarily driven by extra-grammatical factors. A case of this type involves so-called *that*-omission, or complementizer drop, in English, which Dayal & Grimshaw (2009) argues constitutes a type of MCP. However, both experimental (Ferreira & Dell 2000) and modeling (Roland, Elman & Ferreira 2006; Jaeger 2010) work shows that the choice of variant is largely predictable from processing factors.

Moreover, whatever the interpretive properties associated with EV2 turn out to be, it seems clear that they are not identical to those associated with certain other MCP (see examples in (3)). From the point of view of the syntax-meaning interface, the bigger empirical question is whether all MCP share the same interpretive (or distributional) properties, apart from their restricted occurrence in embedded environments. Studying EV2 in the context of theoretical and empirical claims about MCP is therefore important, as it brings us closer to answering the question of what it means to be an MCP, and what the unifying property is, if any. The same can also be said for studying Swedish EV2 in the context of theoretical and empirical claims about EV2 across languages; at the level of the interface of structure and meaning, is EV2 a unified phenomenon? The availability of large-scale naturally occurring data, makes Swedish EV2 particularly well-suited to address these different questions for a type of construction that is both marked and infrequent in speech.

A popular approach to EV2, going back to Hooper & Thompson's now classical work, is to argue that embedded clauses with V2 order are *asserted*—although little consensus has been reached in the literature with respect to what it means for a sentence to be asserted (Grice 1957; Stalnaker 1974; 1978; et seq), or what specific notion of assertion is relevant to the licensing of EV2 (e.g. Andersson 1975; Green 1976; Wechsler 1991; Holmberg & Platzack 1995; Truckenbrodt 2006; Julien 2009; Wiklund 2010; Gärtner & Michaelis 2010; Jensen & Christensen 2013; Julien 2015; Woods 2016a; b). On the other hand, it has been argued that what actually matters for the licensing of EV2, and embedded MCP more generally, are lexical properties of the matrix predicate. However, no consensus has yet been reached regarding the type of lexical properties that determine or constrain the distribution of embedded MCP/V2 (see for instance Den Besten 1983; Weerman & de Haan 1986; Iatridou & Kroch 1992; Vikner 1995; De Haan 2001; Bentzen et al. 2007; Wiklund et al. 2009; De Cuba & Ürögdi 2009; 2010; Haegeman & Ürögdi 2010; Haegeman 2014; Kastner 2015). In this paper we argue, based on statistical data extracted from a series of large-scale written Swedish corpora, that the semantic-pragmatic notion driving the distribution of EV2 is **discourse novelty**; whether the embedded proposition is treated as discourse-old or new information. While this is fundamentally

a pragmatic notion, it is nevertheless tightly constrained by lexical-semantic properties of the matrix predicate.

We additionally demonstrate the use of diagnostics to differentiate the various underlying factors which drive syntactic optionality. This is to highlight how the type of usage data presented in this paper can indeed inform our understanding of traditional grammatical representations rather than supplanting them. We argue that not all probabilistic output is a reflection of learned gradient cognitive representations; the type of usage data which has popularly been analyzed as resulting from either gradient underlying structure (Bybee 2006; Bresnan 2007) or psycholinguistic factors (Jaeger 2010) can (at least in this instance) be better understood as reflecting categorical grammatical representations and their interaction with discourse context.

The following sub-sections (§1.2–1.4) provide the theoretical and experimental background. Section 2.1 details the methods of our study. In Section 2.2 we consider a number of potentially relevant usage- or processing-based factors (Ferreira & Dell 2000; Bybee 2006; Jaeger 2010), which while not previously applied to EV2, nonetheless make testable predictions in this case. We show that, while *stylistic* factors such as formality play a clear role in conditioning the rates of EV2 across grammatical contexts, we do not find evidence supporting a processing or usage-based account. In Section 3, we discuss and test the predictions made by previous influential accounts regarding the type of lexical factors that influence the distribution of EV2. We find that, while certain aspects of the predictions made by these lexical licensing, or selection-based accounts are borne out, as they stand, these accounts are in themselves unable to account for the overall patterns in the data. Section 4 motivates and develops our theoretical account, whereby EV2 is licensed by discourse novelty. We show that this account makes novel predictions about the interaction of certain clause-embedding attitude verbs with matrix negation regarding the availability of EV2. Section 5 presents new experimental data from an acceptability judgment task, showing that these predictions are indeed borne out. Section 6 connects the experimental results back to related corpus data, providing further evidence in favor of our account. Section 7 concludes.

1.2 Main Clause Phenomena

Adding to the observation made by Emonds (1970), that certain types of syntactic structures appear to be confined to matrix clauses, Hooper & Thompson (1973) [H&T] argued that, additionally, certain *classes* of predicates (1), but not others (2), also allow for these structures in their complements.

(1)   Predicate types that allow MCP:
      a.   (Non-factive) speech act predicates, e.g. *say, argue, tell, claim* (H&T's Class A)
      b.   Doxastic non-factives, e.g. *think, guess, believe, imagine* (H&T's Class B)
      c.   Doxastic factives (also known as "semifactives", following Karttunen 1971), e.g. *find out, realize, discover, be aware* (H&T's Class E)

(2)   Predicate types that do not allow MCP:
      a.   (Non-factive) response predicates, e.g. *deny, doubt, accept, admit* (H&T's Class C)
      b.   Emotive factives, e.g. *regret, appreciate, resent, be glad* (H&T's Class D)

Classic examples of English MCP include VP-preposing (3a), topicalization (3b), left/right-dislocation (3c).

(3)　　　Hooper & Thompson (1973: 467–8)
　　　a.　i. Mary plans for John to marry her, and [marry her]$_i$ he will t$_i$.
　　　b.　i. [Each part]$_i$ Steve examined t$_i$ carefully.
　　　c.　i. [This book]$_i$, it$_i$ has the recipe in it.
　　　　　ii. You should go to see it$_i$, [that movie]$_i$.

The observation made by H&T, illustrated in (4) using VP-preposing, is that MCP appear to be possible in the complements of the predicates in (1), but not embedded under those in (2).

(4)　　　Mary plans for John to marry her, and…
　　　a.　I {say, think, know} that [marry her]$_i$ he will t$_i$.
　　　b.　*I {resent, deny} that [marry her]$_i$ he will t$_i$.

For EV2 declaratives,[1] the received view is that V2 is possible under the predicate classes in (1), but not under those in (2), as shown in (5). V-in situ, on the other hand, is the unmarked option, possible under all of the five predicate types in (1) and (2), as shown in (6). (V2 is diagnosed here by V<Neg order, whereas V-in situ is identified by Neg<V order; see Section 1.3.).

(5)　　　*Swedish*
　　　a.　Jon {sa/trodde/visste}　　att　han **hade** inte sett　filmen.
　　　　　Jon {said/thought/knew} that he　had　not　seen movie.DEF
　　　　　'Jon {said/thought/knew} that he hadn't seen the movie.'

　　　b.　*Jon {förnekade/ångrade} att　han **hade** inte sett　filmen.
　　　　　Jon {denied/regretted}　　that he　had　not　seen movie.DEF
　　　　　'Jon {denied/regretted} that he hadn't seen the movie.'

(6)　　　*Swedish*
　　　a.　Jon {sa/trodde/visste}　　att　han inte **hade** sett　filmen.
　　　　　Jon {said/thought/knew} that he　not　had　seen movie.DEF
　　　　　'Jon {said/thought/knew} that he hadn't seen the movie.'

　　　b.　Jon {förnekade/ångrade} att　han inte **hade** sett　filmen.
　　　　　Jon {denied/regretted}　　that he　not　had　seen movie.DEF
　　　　　'Jon {denied/regretted} that he hadn't seen the movie.'

It's worth noting, however, that these empirical claims are based on subtle judgments about the acceptability of the relevant sentences, and that their empirical status is still a matter of debate. For instance, regarding the availability of topicalization in English embedded declaratives, Bianchi & Frascarelli (2009) provide examples like that in (7), which was judged to be acceptable by 80% of their consultants (12/15), thus casting some doubt on the claim that emotive factives disallow MCP.

(7)　　　Bianchi & Frascarelli (2009: 69)
　　　I am glad that [this unrewarding job]$_i$, she has finally decided to give up t$_i$.

---

[1] V2 is also possible in other types of embedded environments, including certain types of adverbial clauses. Here, we leave these to the side, but see for instance Wechsler (1991) and Heycock & Wallenberg (2013).

Here we probe this question in the context of Swedish EV2, which we briefly introduce in the following section.

1.3 Swedish embedded Verb Second

Syntactically, EV2 in Swedish involves movement of the finite verb to C.[2] Importantly, V-to-C languages are different from V-to-T languages. In the latter type, unlike in Swedish and other V-to-C languages, V_{fin}<Neg order is obligatory in *all* tensed matrix and embedded clauses.[3]

In Swedish, which is SVO, it is not always clear from the surface constituent order whether a subject-initial clause has undergone V-to-C movement or not. This is because such movement often results in the same surface-order as a clause without movement, as shown in (8).[4]

(8)   *Swedish*
      a.   Hon gillar katter.
            she likes cats
            'She likes cats.'
      b.
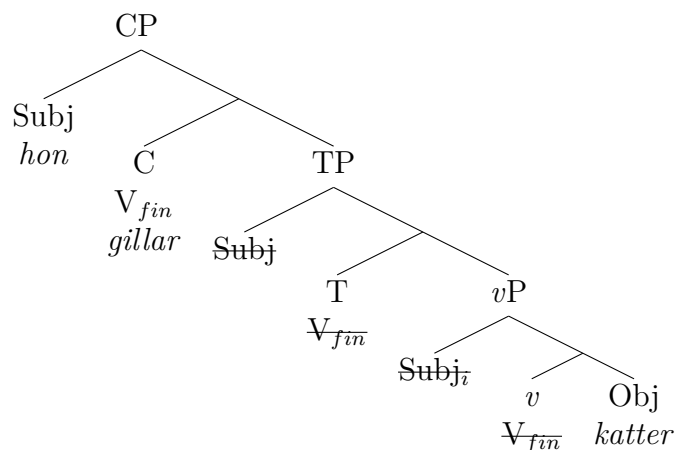
In Swedish, there are two common diagnostics for identifying verb movement. The first is the presence of sentence adverb (including negation), occupying the left edge of vP, as shown in (9b). The second is the the presence of a topicalized or focused non-subject XP in Spec,CP (10).
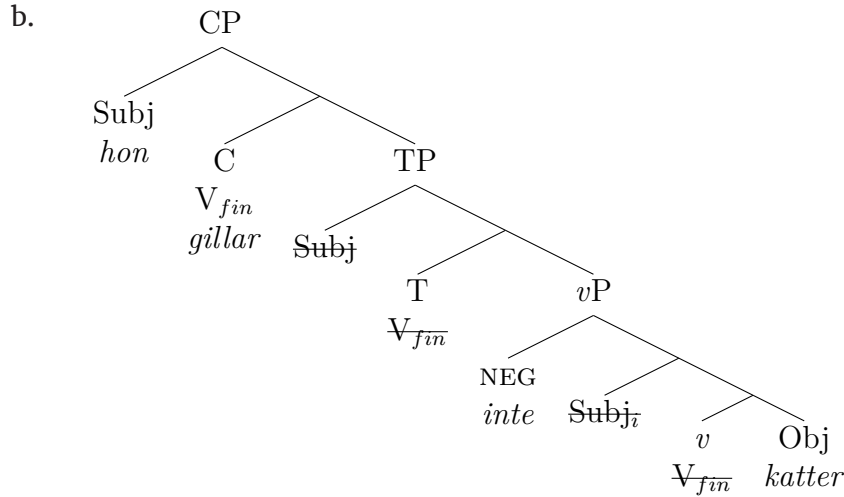
(9)   *Swedish*
      a.   Hon gillar inte katter.
            she likes not cats
            'She doesn't like cats.'

---

[2] Note that this description is likely somewhat simplified. Following Rizzi (1997), much work on V2 and related phenomena has argued that the C-domain consists of an ordered sequence of syntactic heads; Force, Fin, Topic, Focus, etc, responsible for different types of movement-phenomena (see Section 1.4). Since our focus here is the licensing conditions on EV2, and not its precise syntactic implementation, we use C here, as a descriptively simpler and more theory-neutral label.

[3] See for instance Platzack (1987); Platzack & Holmberg (1989); Holmberg & Platzack (1991; 1995); Holmberg (2015), for differences between the different Scandinavian languages in this respect. see also Holmberg (2015) for a recent survey of V2-phenomena.

[4] This is unlike in V2-languages that are SOV, like German and Dutch.

b.

As shown using these diagnostics in (10) and (11), V2 is obligatory in Swedish matrix clauses.

(10)   *Swedish*
    a.   [Den filmen]$_i$ **gillade** hon t$_i$.
         [that movie.DEF]$_i$ liked   she
         'That movie, she liked.'   EV2
    b. *[Den filmen]$_i$ hon **gillade** t$_i$.
         [that movie.DEF]$_i$ she   liked
         'That movie, she liked.'   *V-in situ

(11)   *Swedish*
    a.   Jon **hade inte** sett   filmen.
         Jon had   not   seen movie.DEF
         'Jon hadn't seen the movie.'   EV2
    b. *Jon **inte hade** sett   filmen.
         Jon not   had   seen movie.DEF
         'Jon hadn't seen the movie.'   *V-in situ

While EV2 is possible in certain embedded contexts, as shown in (5)–(6), it is by no means *obligatory* in these contexts, as shown in (12).

(12)   *Swedish*
    Jon {sa/trodde/visste}   att han (**hade**) **inte** (**hade**) sett   filmen.
    Jon {said/thought/knew} that he   had   not had   seen movie.DEF
    'Jon {said/thought/knew} that he hadn't seen the movie.'

Next, we turn our focus to the interpretive effects typically associated with EV2, and MCP more broadly.

1.4 Interpreting MCP

The received view in the literature going back to Hooper & Thompson (1973) is that Main Clause Phenomena are associated with **illocutionary force**, the type of speech act associated with an utterance. For declaratives, *assertion* is the associated illocutionary force. Following Stalnaker (1974), a speaker asserting a proposition p minimally requires that:

(13) a. The speaker is committed to p;
      b. The speaker is attempting to add p to the Common Ground [CG] (the set of propositions mutually taken to be true by the discourse participants).

It is uncontroversial that in uttering either sentence in (14), the speaker is typically asserting something about their beliefs, and not about John.

(14) a. I believe the rumor about John.
      b. I believe that John stole the money.

There does, however, exist a reading of (14b) on which the speaker is asserting the proposition that John stole the money. On this reading, the matrix clause "I believe…" plays a parenthetical role. As was observed already by H&T, the latter reading can be paraphrased using a slifting construction, as in (15).

(15) John stole the money, I believe.

Connecting the availability of MCP to the presence of illocutionary force would then nicely capture both their obligatory occurrence in matrix clauses, as well as their restricted availability in embedded clauses.
   One popular way of encoding this connection between the syntax and the pragmatics is to say that MCP involve an extended C-domain that encodes illocutionary force (such as that in (16) from Rizzi 1997), as well as other discourse features like topic and focus. V-to-C movement is then argued to be triggered by interpretable features on Force.

(16) Rizzi (1997: 297)
      $[_{ForceP}$ Force $[_{TopP}$ Top $[_{FocP}$ Foc $[_{TopP}$ Top $[_{FinP}$ Fin IP $]]]]]$

This can be contrasted with clauses that disallow MCP, which involve a smaller, or "impoverished" C-domain, in (17), incompatible with illocutionary force, topicalization and focus, as well as with any movement to their dedicated positions in the left-periphery, including V-to-C.

(17) $[_{FinP}$ Fin IP $]$

A problem for this perspective arises, however, when we consider factive predicates like *discover* or *realize*.

(18) John discovered that $[_p$ Anna likes cats].

On the classic view of assertion, given in (13), factive predicates are predicted to disallow embedded assertions, given that factives *presuppose* that p is true (Kiparsky & Kiparsky 1970; Keenan 1971; Karttunen 1971; 1974). This is because—on the received view of presupposition (Stalnaker 1974; Heim 1982; 1983; 1992)—for a sentence involving a factive predicate to be felicitous, p must be entailed by the context (the intersection of the propositions in the Common Ground; i.e., the worlds in which all of the propositions in the Common Ground are true). That is, p must already be part of the Common Ground. On this view then, factivity is incompatible with the second component of assertion, in (13b); the speaker attempting to *add* p to the Common Ground. As we saw above, however, V2 and other MCP have been observed to be possible in clauses embedded under at least the doxastic factives.

In contrast, some authors (e.g. De Cuba & Ürögdi 2009; 2010; Haegeman 2010; Haegeman & Ürögdi 2010; Haegeman 2012, and Kastner 2015), have nevertheless claimed—as would be expected given the standard views of factivity and assertion, respectively—that factive verbs as a class disallow MCP. This has been supported by data points such as the following:[5]

(19)   a.   Maki, Kaiser & Ochi (1999: 3)[6]
              *John regrets that this book Mary read.

        b.   Hegarty (1992: 52; fn. 19)[7]
              *Mary realizes that this book, John read.

While the above authors take this to be a general empirical claim about MCP, it has gained less traction in the literature on EV2. To our awareness, the only authors to advance this claim are Truckenbrodt (2006) and Reis (1997) in the context of German EV2.[8] It is nevertheless important to consider this view seriously for Swedish EV2, both in view of the general question of whether MCP constitute a homogeneous class, and in terms of the more specific question of the type of predicates that allow EV2, specifically in light of the conflicting nature of some of the empirical claims made regarding the distribution of MCP. We return to the question about the role of factivity, and address it empirically in Section 3.1.[9]

In accounting for their observation that the doxastic factives seem to allow MCP, Hooper & Thompson (1973: 481) claim that while these verbs are presuppositional in the traditional sense, the doxastic factives (like the speech act and doxastic non-factives), "have a parenthetical reading on which the complement proposition is considered the main assertion." More recently, this idea has been taken up by Jensen & Christensen (2013) in the context of EV2. Rather than using the already theory-laden label of "assertion", these authors have adopted the notion of the Main Point of the Utterance from Simons (2007). This is (roughly) the content of an utterance which most directly addresses the Question Under Discussion (Roberts 1996; 2012). As illustrated in (20), given the question in (20-Q), the Main Point of Utterance in both (20-A1) and (20-A2) is *John is in New York*.

(20)   Q.   Where is John?
        A1.   [$_p$ He's in New York.]
        A2.   I think that [$_p$ he's in New York].

As observed by Simons (2007), doxastic factives, like the speech act and non-factive doxastic predicates—but apparently unlike the emotive factives—allow the embedded clause to provide the Main Point of the Utterance.

---

[5] Although see (7) from Bianchi & Frascarelli (2009) above for conflicting empirical claims about the status of topicalization in clauses embedded under emotive factives.

[6] Also cited in De Cuba & Ürögdi (2010: 43), Haegeman & Ürögdi (2010: 112), Haegeman (2012: 257), Kastner (2015: 3), De Cuba (2017: 4).

[7] Also cited in Haegeman & Ürögdi (2010: 113), Haegeman (2012: 257), De Cuba (2017: 4).

[8] Though Truckenbrodt (2006: 299) nevertheless reports "exceptions" to this generalization.

[9] There are other interesting and important components to the analyses presented here; for instance, these authors link the "presuppositional" nature of factive clauses to the selection of a "referential" CP (following Kiparsky & Kiparsky 1970, some authors argue that this is encoded as an overt DP; see also Adams 1985; Pesetsky 1991; Rooryck 1992; Bhatt 2010; Abrusán 2011; Elliott 2016; Bogal-Allbritten & Moulton 2018; Djärv 2019a; b, for discussion). What is important on these accounts, is actually the status of the embedded clause as referential or presuppositional. However, what matters for current purposes, is that these authors assume that all factive predicates are presuppositional in the relevant sense, and therefore predicted to disallow EV2 and other MCP.

(21) Q. Where is John?
      A. I found out that [$_p$ he's in New York].
      A. #I'm happy that [$_p$ he's in New York].

The claim advanced by Jensen & Christensen (2013) is that it is this notion of Main Point content that distinguishes between those embedding environments that allow MCP, and those that do not. On this view, any observed predicate restriction on EV2 is essentially epiphenomenal, reflecting simply the relative ease with which a given predicate may function parenthetically.

  This account, however, is problematic for purely empirical reasons. For instance, Wiklund et al. (2009) present judgment data showing that neither is V2 obligatory in these contexts, nor is it ruled out in a context where the embedded proposition is not the Main Point of the Utterance, as illustrated with the question-answer pair in (22).

(22) *Swedish* (Wiklund et al. 2009: 1929)
      a. Varför kom han inte på festen?
          why   came he   not   to party.DEF
          'Why didn't he come to the party?'

      b. Kristine sa   att han **fick**   inte.
          Kristine said that he   was.allowed not
          'Kristine said that he wasn't allowed to.'   EV2

According to Wiklund et al. (2009), this sentence, in the context of (22-Q), can either be read as 'he didn't come to the party because he wasn't allowed to, as Kristine told me', or 'he didn't come to the party because Kristine said that he wasn't allowed to go'. However, on a strong version of this hypothesis, the second reading should not be available. On the alternative view advanced by Wiklund et al. (2009), a predicate will allow V2 in its complement if it *also* allows for the embedded proposition to be the Main Point of the Utterance, which, unlike previous authors, they take to be a matter of selection. Crucially then, there is no *direct* link between assertion and EV2. Rather, the predicates in (1) select for a larger CP, like that in (16), which is *compatible* with V-to-C movement, *as well as* with the illocutionary force of assertion. The predicates in (2) however, select for a smaller CP, as in (17), which they take to be compatible with neither V-to-C, nor with illocutionary force.

  However, noting that the critical judgments are subtle and based on the intuitions of only a few speakers, Djärv, Heycock & Rohde (2017) tested experimentally whether participants' judgments of acceptability for sentences with EV2 in Swedish were sensitive to this type of context manipulation. The form of the manipulation they used is illustrated with the English examples in (23). The prediction was that EV2 should be acceptable in contexts like (23a), but not in contexts like (23b).

(23) *Swedish* (Djärv et al. 2017: 6)
      a. Q. Why didn't Kate come to the party?   [Main Point: EC]
          A. John thinks that [$_p$ **she's left town**].

      b. Q. Why didn't John invite Kate to the party?   [Main Point: MC]
          A. **John thinks that** [$_p$ **she's left town**].

Their experiment, which was a judgment study, manipulated Main Point status [matrix clause; embedded clause], predicate type [Speech Act; Doxastic Non-factive; Doxastic Factive; Emotive Factive], and word order (V<Neg; Neg<V). They found a main effect of word order such that V3 (subject-adverb-verb order) was rated overall higher than V2 ($p < 0.001$),

as well as a significant effect of predicate type ($p < 0.001$): speech act predicates and doxastic factives were rated higher than the doxastic non-factives and the emotive factives (in line with corpus results from Danish cited in Jensen & Christensen 2013). However, there was neither a main effect of Main Point status ($p = 0.88$), nor was there an interaction with Main Point status ($p > 0.75$), contrary to an account where V2 is driven by the asserted status of the embedded proposition.

   These results are problematic for the view that MCP and EV2 are driven by the Main Point status of the embedded proposition. Rather, their results seem more in line with a lexical licensing account, whereby the acceptability of EV2 is driven purely by the type of embedding predicate, such as that advanced by Wiklund et al. (2009). Here, EV2 is licensed only by certain predicates, and is not associated with a particular discourse status. Although this account appears to correctly capture the pattern of data seen above, this type of account leaves open the question of exactly what distinguishes the cases where V2 does occur and when it does not. That is, if we adopt this view, we seem to be forced to adopt the view that EV2 is truly optional.

   Finally, recall the first component of assertion, given in (13a); that the speaker is committed to p. We noted above that the Common Ground component of assertion, in (13b), is problematic given factive predicates. A number of authors, however, have argued that what is relevant for the licensing of EV2 is in fact only the criterion in (13a). Truckenbrodt (2006), building on Wechsler (1991), argues in the context of German V2 that EV2 is possible as long as someone in the context (either the speaker or the matrix clause subject) believes that p is true.[10] (See also Wiklund 2010; Julien 2015; Woods 2016a; b for different versions of this general perspective.) Evidence against such a view, however, comes from Gärtner & Michaelis (2010), looking at German V2. They observe that although in a sentence involving matrix clause disjunction, the speaker is committed to neither of the disjuncts, V2 is nevertheless well-formed (and in fact obligatory!) in such sentences:

(24)   *German* (Gärtner & Michaelis 2010: 4)
         In Berlin **schneit** es oder in Potsdam **scheint** die Sonne.
         in Berlin snows   it or   in Potsdam shines   the sun
         'It is snowing in Berlin or the sun is shining in Potsdam.'

The same is also true in Swedish:

(25)   *Swedish*
         Antingen **snöar** det i Umeå, eller så **skiner** solen   i Skellefteå.
         either   snows it in Umeå or   so shines sun.DEF in Skellefteå
         'It is ether snowing in Umeå or the sun is shining in Skellefteå.'

Gärtner & Michaelis (2010) present a view according to which V2 involves a weaker notion of assertion than that given in (13); rather than operating at the level of speech acts, they take the relevant notion of context update to be one which operates only at the propositional level. Their analysis of a sentence like (24) is given in (26):

(26) $[[\text{p-V2 or q-V2}]] = [\, p \cap CG \,] \cup [\, q \cap CG \,]$

Noting however, that their account nevertheless over-generates, in the case of matrix negation and conditionals, neither of which allow V2, they add a so called "progressivity requirement on assertive update":

---

[10] According to him, matrix clause V2 additionally requires that the speaker wishes to add p to the Common Ground.

> **Progressive update** (Gärtner & Michaelis 2010: 9)
> "An assertive update CG' of a common ground CG by an utterance u_d containing meaning components $\blacktriangleright\phi_1 \dots \blacktriangleright\phi_n$ is *progressive* if CG' $\subseteq$ *eq* [CG $\cap$ $(\phi_1 \cup \dots \cup \phi_n)$]."

They further state that "Progressive update captures the intuition that (dependent) root phenomena in general, and V2-declaratives in particular, come with an informativity requirement related to providing "new information"" (Gärtner & Michaelis 2010: 10).

In this section, we discussed different accounts of the type of interpretative effects associated with EV2 (and MCP more broadly). We noted that previously reported (experimental and judgment) data on Swedish EV2 only appears to be compatible with an account, such as that in Wiklund et al. (2009), whereby EV2 is licensed, but not obligatory, under certain predicate classes ((1), (2)). However, we noted that these judgments about EV2 (and MCP more generally) are subtle and appear to vary across speakers, both in terms of the acceptability of EV2 and any proposed associated semantic-pragmatic correlates. Given the current state of the literature, it is possible that there are other factors beyond the type of embedding predicate—either grammatical, contextual, or processing-based—that may influence the use and distribution of the two variants. The remainder of this paper sets to tease such potential factors apart.

Section 2.1 details the methodology for extracting corpus data. Section 2.2 is devoted to testing various processing-based hypotheses, and Section 3 to testing the predictions made by the two types of lexical licensing accounts discussed above. We show that the actual patterns of usage are not compatible with either of these accounts. Neither are they compatible with a processing or usage-based account of EV2. From considering the types of embedded environments in which V2 is licensed, we arrive at a pragmatic licensing account, whereby EV2 is licensed by discourse novelty. This view then, ends up being entirely compatible with that proposed by Gärtner & Michaelis (2010), developed to account for the distribution of main clause V2 in German.[11] The following section details the methods of the corpus study.

2 Manifestations of grammar in usage

2.1 Corpus methods

We extracted natural language usage data from several very large Swedish corpora (Borin, Forsberg & Roxendal 2012) totaling 12,873,778 sentences, subsequently referred to as BFR (from the authors Borin, Forsberg and Roxendal). BFR also represents a balanced set of genres ranging from informal blogs and forums to formal academic writing and government texts. These are summarized in Table 1.

Owing to the Zipfian distribution of frequencies inherent to language use (Yang 2013; Piantadosi 2014), the majority of sentences only include a limited number of highly frequent verb types, with most predicates occurring only rarely. As such, the large sample of extracted data is required for the type of analysis presented in this paper. This is particularly relevant since we find that only about 5% to 10% of sentences provide a diagnostic

---

[11] Although note that it's less clear how their account would deal with V2 in German *neither, nor…* sentences, such as:

(1) *German*
     Weder schneit es in Berlin, noch scheint die Sonne in Potsdam.
     neither snows it in Berlin, nor shines the sun in Potsdam
     'It's neither snowing in Berlin, nor is the sun shining in Potsdam.'

It appears to us that such sentences, which would presumably be interpreted as $\neg(\phi_1 \vee \phi_2)$, are at odds with their progressive update criterion for EV2. We leave this issue to the side. Thanks to Florian Schwarz, p.c. for this observation.

Table 1: Rates of EV2 across corpora of varying formality. "Genre" represents a coarse categorization of corpora by source material. "Corpus" is the division provided within BFR. "Sentences" is the total number of sentences extracted from the original sub-corpus. "Proportion Non-ambiguous" represents the proportion of sentences within each subcorpus over which our extraction algorithm is able to apply the diagnostic for estimating EV2 vs. in-situ status. "p(ev2)" is the proportion of such sentences surfacing with EV2 order rather than embedded in-situ. Note that while the proportion of diagnostic cases is more or less steady by corpus, there is a clear effect of genre on the rates of EV2. Formal or more heavily prescriptive content has lower rates of EV2 compared to colloquial and informal material. Even in the most formal styles EV2 is still consistently attested.

Genre	Corpus	Sentences	Proportion Non-ambiguous	p(ev2)
Blogs and Forums	Familjeliv-känsliga	5971907	0.1163	**0.0636**
	Familjeliv-nöje	458699	0.0809	**0.0555**
	Familjeliv-adoption	77008	0.0936	**0.0545**
	Familjeliv-expert	57478	0.0966	**0.0522**
	Bloggmix	2713376	0.0765	**0.0502**
	Flashback-Politik	2841872	0.0972	**0.0457**
Historical	Tidning 1870	17084	0.06	**0.0724**
	Tidning 1860	58839	0.062	**0.0512**
Academic	Sweacsam	52678	0.0736	**0.0375**
	Humanities	60931	0.0741	**0.0283**
Goverment	Rd-bet	372054	0.0698	**0.0163**
	Rd-ds	172657	0.0848	**0.0141**
	Rd-fpm	5259	0.0686	**0.0138**
	Rd-skfr	81800	0.0865	**0.0098**
Accessible	Åttasidor	8059	0.0768	**0.0081**

test of EV2 status,[12] and of those, EV2 order is only used approximately 5% of the time. This means that one would need to analyze on the order of 40,000 sentences to encounter 100 diagnosably positive examples.

As the goal is to examine sentences with the potential for EV2 order (regardless of whether or not that was actually realized), we created a subcorpus for analysis according to the following method.[13] Data was collapsed according to the lemma tags which were automatically assigned in BFR. The use of *lemma* here does not reflect a theoretical assumption regarding underlying roots, but is simply a limited technical implementation aimed at providing a single representation across surface-divergent inflected forms.[14] The analysis was also replicated over raw inflected verb forms and we did not identify any major qualitative differences. However, the use of lemmas reduces data sparsity; even in a large corpus many possible inflected forms are unattested, and so grouping together inflectional variants can alleviate that. BFR data are not parsed and automatic syntactic parsing faces numerous technical limitations on data of this diverse type and size (Sekine 1997; McClosky, Charniak & Johnson 2010). Instead, we utilized several filters over BFR-provided part-of-speech tags

---

[12] This estimate comes from the present study; the proportion of sentences with an overt adverb such as negation in the embedded clause.

[13] Code is available open-source at https://github.com/scaplan/ev2-optionality under the MIT license for replicability and extension to related data sets and analyses.

[14] For example, *spring, springa, springer, sprang,* are all identified by the unifying lemma *spring* ('run') and identified as such in the subsequent analysis.

(Brill 2000) in order to differentiate cases in which an embedded verb has remained in situ rather than undergone V-to-C movement.

For technical simplicity, we only consider single, rather than multiple, embeddings (approximately 20% of all sentences with the overt complementizer *att* contain more than one instance). Sentences are further excluded if the complementizer is directly followed by a verb, with no intervening potential subject information, since this is indicative of a non-finite complement rather than a tensed embedded clause. Additionally, we exclude sentences in which the matrix verb is the copula, since these can correspond to a broad range of predicate types. A few additional filters exclude potential false-positives such as future-marking *kommer att* ('will'), adverbial clauses involving *eftersom/(där)för att* ('because'), and embedded clauses with relative clause subjects, as these are problematic for unambiguously identifying the tensed verb of the embedded clause.

This set of embedded complement sentences is diagnosed for EV2 status by considering the relative linear order of the embedded verb and negation (as outlined in Section 1.3). Theoretically, this diagnostic can be applied with any adverb in the embedded clause, however for tractability we limit our diagnostics to negation (*inte, icke*, etc.).

This results in a set of EV2/in-situ sentences which is necessarily a subset of the total instances in the data. However, there is no theoretical reason to expect factors such as non-negation adverbials or multiple embedding to have a profound and significant impact on the realization of EV2. Limiting our search to single-embedded sentences with negation allows technical tractability and high-confidence in the quality of output data while still providing a representative sample of over one million diagnosed sentences.

A range of statistical information was additionally extracted for each sentence and for each lemma overall. This includes frequencies, lexical semantic information such as [H&T] class (see Section 1.2),[15] polarity information, and several conditional probability events (e.g. matrix introducing embedded clause, matrix introducing EV2 clause, embedded predicate surfacing in embedded clause, embedded predicate surfacing with EV2 order, etc.) A full enumeration of extracted information is available in the source code.

2.2 Lexical information and variation

At a descriptive level, Table 1 provides a summary of EV2 by corpus. We find that overall rates of EV2 are graded by formality, with more colloquial Swedish such as blog and forum text exhibiting higher rates than formal writing. This is consistent with Heycock & Wallenberg (2013) who find comparable rates in blogs and spoken caregiver data compared with novels. This potentially reflects a sociolinguistic property of a prescription against EV2. It is striking, though, that EV2 appears stable diachronically, without a significant change between historical newspaper texts dating back to the 1860's and modern online forums.[16] This stability suggests that synchronic proportions of use do not represent a case of language change in progress, but rather a fact about the interaction of grammatical representation and use in context. The majority of our subsequent analyses were conducted primarily on the Flashback-politik subset of our data. This was done since it relates more closely to spoken dialogue compared with some of the other written material.

If we examine the attested likelihood of EV2 order by matrix predicate, we notice a fairly large degree of variation (Table 2). While the overall rate of EV2 varies by genre (see Table 1), there is general cross-corpus consistency in the *relative* rates of EV2 for

---

[15] A highly frequent but limited set of 108 lemmas was tagged for semantic class based on their classification in previous literature on the topic (Hooper & Thompson 1973; Wiklund et al. 2009; Kastner 2015; Djärv et al. 2017).

[16] As in Table 1 the rates of EV2 in historical newspaper data (1860–1880) range from 5% to 7% in line with the contemporary rate in online forums.

Table 2: Sample values of probability of EV2 by predicate. Data from the Flashback-Politik corpus.

Verb	Gloss	p(ev2)
drömma	'dream'	0.00
glömma	'forget'	0.12
höra	'hear'	0.08
säga	'say'	0.07
tro	'believe'	0.03
tycka	'think'	0.06

individual verbs. This raises an important question of how to understand language usage data: Should we take any observed rate of use to be *definitional*? In other words, is the fact that *glömma* ('forget') introduces EV2 at four times the rate of *tro* ('believe') acquired from direct input by learners and subsequently recapitulated for the next generation? Or do rates of EV2 emerge in some other capacity? More generally, what can we learn by different ways of modeling usage data?

Previous accounts of similar data (Bresnan et al. 2007) have attempted to demonstrate that usage patterns are predictable, and thus worth studying on top of traditional grammatically judgment data. Yet, we should note that a statistical model which predicts linguistic output with high accuracy is not, in and of itself, an explanatory theory. It is insufficient to model the outcome of syntactic optionality unless we can move towards understanding *why* correlated variables have predictive power. Statistical tools can only verify, but not produce, empirical hypotheses.

We start from the premise that usage statistics need not be the thorn in the side of generative syntactic and semantic theory, but rather an informative window onto the underlying representations; thus flipping an argument typically taken by usage-based linguists (see Du Bois 1985; Hopper 1987; Bybee 2006, a.o.). Rather than advance the claim that "discourse use shapes grammar" (under which rates of use serve as a direct proxy for grammatical representation), the EV2 alternation presents a case study in how grammatical factors can influence rates of use in discourse instead of the inverse. In particular, we are able to evaluate specific grammatical hypotheses quantitatively.

2.3 Usage-based and processing accounts

Under a usage-based framework, grammar is taken to be simply the cognitive organization of one's experience with language. As such, factors like frequency of use of particular constructions would be predicted to have an impact on their representation. Bybee (2006: 714) summarizes the usage-based view as follows: "[Grammar] does not have structure a priori, but rather the apparent structure emerges from the repetition of many local events." This family of accounts is tightly linked to explanations of optionality and usage rates at the psycholinguistic level. Under this psycholinguistic view, "not only do the syntactic privileges of the to-be-produced lemmas affect syntactic structure, but so too can the timing of lemma selection have important effects on the syntactic structure of a sentence." (Ferreira & Dell 2000: 299). The intuition can be captured by imagining that a speaker wants to recount story of the outcome of the race between the tortoise and the hare in Aesop's famous fable using a verb like *defeat*. In principle, this could be encoded through either can active or passive structure. If the word *hare* is significantly faster to activate from memory than *tortoise* then it would be more efficient for the speaker to output that constituent first using the passive (*the hare was defeated by the tortoise*) compared with the active (*the tortoise defeated the hare*). Otherwise the already-retrieved *hare*

would need to sit around in a buffer waiting to be output until the rest of the sentence had been uttered.

This psycholinguistic account has been applied to similar cases of "syntactic optionality" such as *that*-omission[17] (Ferreira & Dell 2000), and makes straightforward predictions with respect to EV2 in Swedish: there is a "race" between outputting the adverb (in this case negation) and the embedded predicate. If the embedded predicate is activated first then it is output before the adverb (EV2 order) whereas if negation is activated first then it is output before the predicate (V-in situ order). On this theory, any factors which correlate with faster word activation will also be proxies for the corresponding output order. Such factors are well-studied empirically and include frequency (Ferreira & Dell 2000; Bock & Levelt 2002) and predictability of syntactic structure (Jaeger 2010; Hale 2014; Caplan 2018).[18]

In light of this, it is worth evaluating the fit of the same factors in the case of EV2. A processing or usage-based account should predict a connection between the frequency with which matrix predicates introduce embedded clauses (since it should speed up access of subsequent required embedded structure) and rates of EV2. However, as is clear in Figure 1, there is no relationship between the probability of a verb introducing an embedded clause and the probability of EV2.
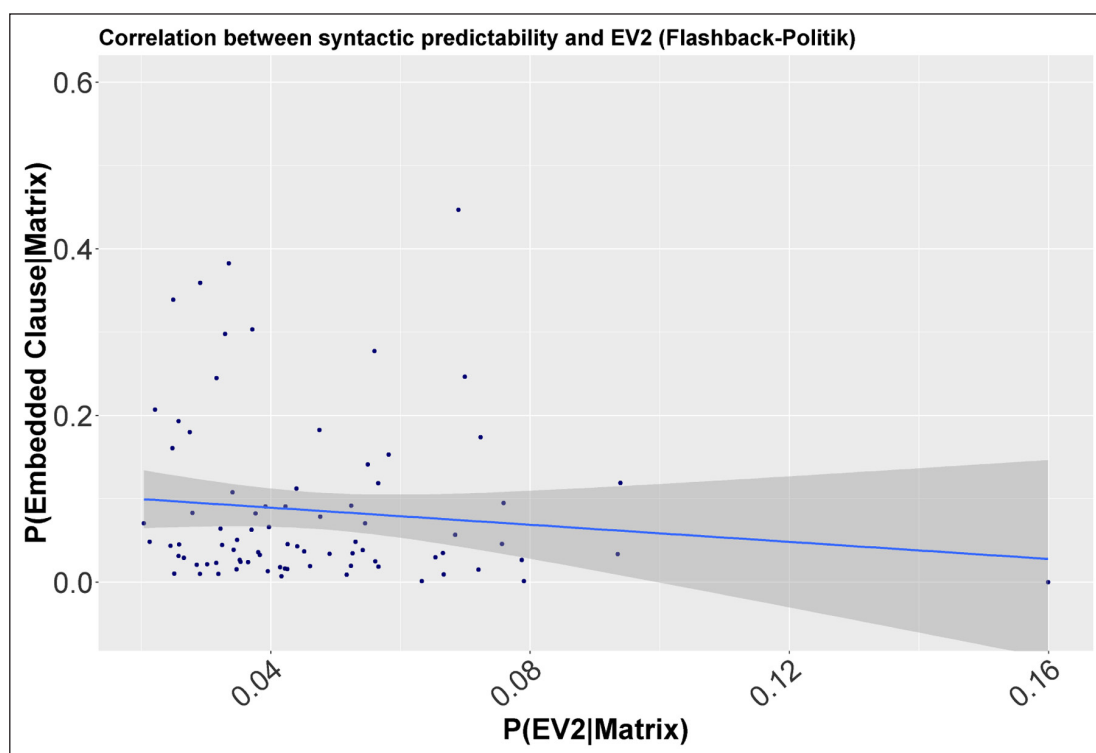
Figure 1: Probability of EV2 (X-axis) against the probability of introducing an embedded clause for each matrix predicate. This is limited to verbs which have a minimum frequency of 1000, introduce an EV2 clause at least 5 times, and which have occurred in a diagnostic sentence at least 100 times in the Flashback-politik corpus. There is no significant correlation between embedded clause-taking and the likelihood of EV2 order.

---

[17] The alternation illustrated in (1a) vs. (1b).

(1) a. The coach said the players were tired.
     b. The coach said that the players were tired.

[18] What's more, the availability of lemmas for fast processing can be experimentally manipulated to causally induce the use of passive structures over otherwise equivalent active ones (Gleitman et al. 2007).

This is confirmed by a linear regression model in which *probability* of EV2 (*conditioned on matrix verb*) is the dependent measure and *probability of embedded clause* (*conditioned on same matrix verb*) is the independent measure (Table 3). A similar prediction might be made within the embedded clauses themselves—there might be some connection between the frequency of a verb appearing in an embedded clause and EV2. This prediction is not borne out; there is no significant correlation between the frequency with which a verb appears in an embedded clause and the rate at which it occurs in an EV2-clause (Table 3). Nor is there any correlation between the total frequency of a verb and the rate at which it occurs in an EV2-clause. Analyses are stable across corpora, but results below are reported only from the Flashback-politik corpus.

It is still conceivable that an underlying relation is hidden by the fact that the majority of lemmas are never attested with EV2. However, this is not the case; these analyses are robust even if limited only to verbs attested as taking EV2 order at least five times (Table 4).

Another possibility is that rates of EV2 are a psycholinguistic by-product of speakers "forgetting" that they're in an embedded clause, something akin to a speech error or disfluency. It is impossible to directly quantify the degree of disfluency based on text alone, but we can take an estimatable proxy. If a large amount of syntactic material or information content intervenes between the matrix verb and the beginning of the embedded clause, the processing system might be more likely to reset to applying main-clause syntax. If this were the case, we would predict an increase in intervening material before the complementizer to correlate with increased rates of EV2. In practice however, there is no clear relation between intervening material and EV2 (Figure 2).

What should be made of this lack of processing-level effects on EV2? The fact that we do not find evidence for a connection between frequency or predictability factors and rates of EV2 order are less a failure to replicate past work (Ferreira & Dell 2000; Bresnan et al. 2007), and more an identification that whatever components drive the realization of EV2 vs. in-situ order are grammatical factors, rather than

Table 3: There is no correlation between likelihood of EV2 order and the probability of the matrix predicate introducing an embedded clause. Nor is there any correlation between likelihood of an embedded verb taking EV2 order and the frequency with which that verb appears in embedded clauses. Analysis is limited to verbs which occurred in a diagnostic sentence at least 100 times in the Flashback-politik corpus.

Dependent	Factor	Estimate	Std. Er.	t-value	P(>|t|)
P(EV2|matrix)	P(EC|matrix)	−0.0126	0.0201	−0.627	0.532
P(EV2|embed)	P(EC|embed)	9.583e-08	2.443e-07	0.392	0.696
P(EV2|embed)	Freq(embed)	−8.922e-09	1.593e-08	−0.560	0.576

Table 4: There is no correlation between likelihood of EV2 order and the probability of the matrix predicate introducing an embedded clause. Nor is there any correlation between likelihood of an embedded verb taking EV2 order and the frequency with which that verb appears in embedded clauses or the frequency of that verb overall. Analysis is limited to verbs which occurred in a diagnostic sentence at least 100 times and occur with EV2 order at least five times in the Flashback-politik corpus.

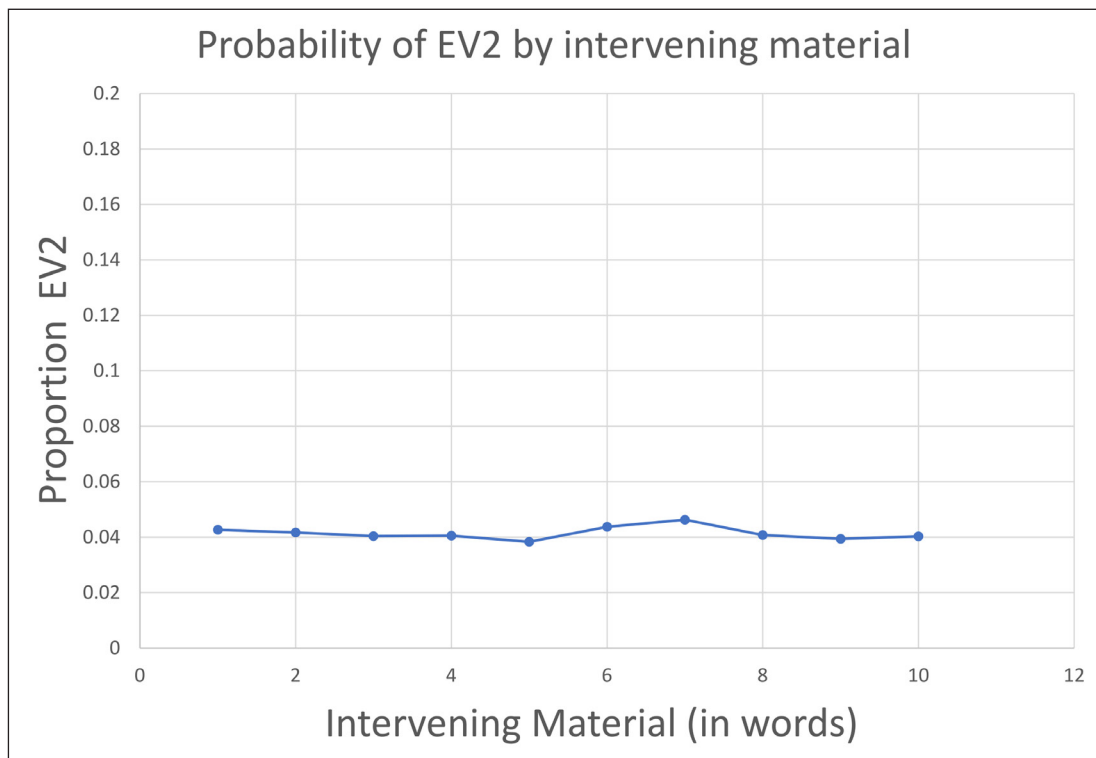| Dependent	Factor	Estimate	Std. Er.	t-value	P(>|t|)
P(EV2|matrix)	P(EC|matrix)	−0.005197	0.017014	−0.305	0.76
P(EV2|embed)	P(EC|embed)	1.950e-08	2.874e-07	0.068	0.946
P(EV2|embed)	Freq(embed)	−1.495e-08	1.760e-08	−0.849	0.397

Figure 2: Length of material (number of words) intervening between the matrix predicate and the complementizer (X-axis) against the rate of EV2 (Y-axis). There is no overall effect of intervening material on rates of EV2, contra the predictions of a sentence production account. Data from the Flashback-Politik corpus.

psychological-production ones. Probabilistic effects are not a homogeneous set and not all "optionality" represents the same kind of phenomenon. While a term like "optional" may imply free and unconditioned variation, research on this topic has found underlying conditioning factors to range from sociolinguistic (Elsness 1984), to morpho-syntactic (Jeoung 2018), to psycholinguistic (Ferreira & Dell 2000). See Tamminga et al. (2016) for an overview of such contrasts. Simply because a sample of language use is probabilistic in nature does not necessarily mean that the underlying linguistic representation is itself probabilistic, contra the usage-based view that rates of particular lexical co-occurrence is also part of the underlying representation (Bybee & Eddington 2006).

We argue that the consistent by-predicate rates of EV2 do not require speakers to be implicitly sensitive to such probabilities, but rather these stable rates of variation emerge as an interaction between meaning and context. A speaker doesn't need an internal counter to tell them to utter a particular syntactic variant (EV2 vs. V-in situ) with *say* twice as often as with *deny*. Rather, the relative rates of EV2 across predicates arise in the interaction between lexical properties of the matrix predicate, the discourse function associated with EV2, and elements of the discourse context. In the following sections, we evaluate several theories about the discourse function of EV2, and the types of lexical properties that are relevant to the licensing of EV2.

3 Lexical accounts of EV2

In this section we test the predictions of the two types of lexical licensing accounts discussed above. First, in Section 3.1, a set of accounts according to which the derivation of MCP (including V2) is *blocked* in certain environments, defined in terms of the presuppositional

requirements of the matrix predicate; specifically under factive predicates. Secondly, in Section 3.2, accounts according to which V2 is available, but entirely optional, under certain predicate types; i.e., those that (independently) license embedded assertions. We test the predictions made by these accounts against BFR data, showing that for neither of these accounts are their predictions straightforwardly borne out.

3.1 Factivity

On the type of account articulated in De Cuba & Ürögdi (2009); Haegeman & Ürögdi (2010); Haegeman (2014); Kastner (2015), among others, factive verbs are predicted to categorically disallow EV2, as a type of Main Clause Phenomena (see discussion in Section 1.4). We noted that this line of analysis is at odds with the observation made by Hooper & Thompson (1973) and subsequent work, that the doxastic factives allow MCP and V2 complements. Nevertheless, given that judgments in this area appear to be subtle and prone to variability, we wanted to test the empirical claim that factive and non-factive predicates differ fundamentally in their ability to license MCP in the context of EV2, against the large scale data available in the BRF-corpora. If these views were correct, we would expect significantly lower rates of EV2 under factive than under non-factive verbs.

However, as shown in Figure 3, we find that factivity does not influence the rates of EV2. In fact, from this plot, it looks as though factive verbs (the gold bar) show slightly higher rates of EV2 than the non-factive verbs (the gray bar); however, this difference is not statistically significant.

We also ran a Wilcoxon Rank Sum test (a non-parametric alternative to the two-sample *t*-test), which allowed us to reject the hypothesis that the distribution of EV2 sentences is different for factive as opposed to non-factive verbs (W = 748, p = 0.6949). This was true for all corpora that we investigated.

3.2 (Optional) lexical licensing

On the view advanced by Wiklund et al. (2009), discussed in Section 1.4, EV2 is optional in the complements of certain predicate types, namely those in (1): speech act predicates, doxastic non-factives, and doxastic factives; but not in the complements of the predicate classes in (2): emotive factives and response predicates.
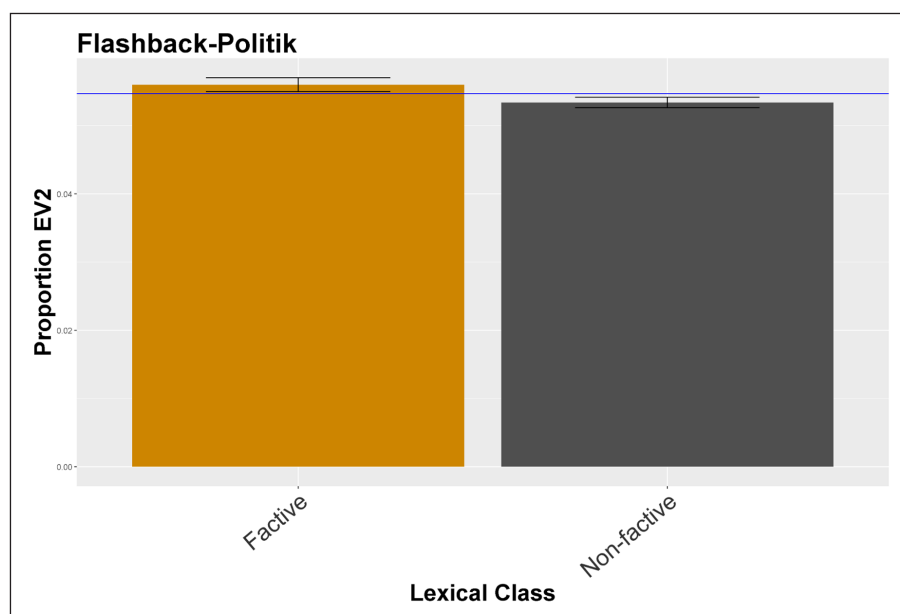

Figure 3: Rates of V2 under factive vs. non-factive verbs; plot based on data from the Flashback-Politik corpus (2,841,872 sentences).

In terms of the distribution of EV2 in the corpus, this account predicts that the relevant factor determining the rates of EV2 is simply membership of a particular lexical class. Moreover, given that pragmatic factors play no explanatory role on this account, we expect that if it were correct, then the rates of EV2 across predicate classes should be essentially constant, both across different discourse types—represented by the different genres of the corpora (see Table 1), as well as across the different predicates within a given predicate class.

Contrary to the first of these two predictions, we find that, while the distribution of EV2 to some extent varies across predicate classes along the lines predicted by this account (overall higher rates of V2 in the complements of speech act predicates, doxastic non-factives, and doxastic factives), the rates of EV2 across predicate classes varied *substantially* across different corpora, as shown in Figure 4.

It is also worth noting that in neither corpus do the rates of EV2 straightforwardly track the rates of EV2 found in Jensen & Christensen's (2013) Danish corpus, which were also reflected in the judgment data from Djärv, Heycock & Rohde (2017), where the speech act and doxastic factives showed the highest rates/judgments of acceptability for EV2, followed by the doxastic non-factives and the emotive factives.

Moreover, contrary to the second prediction made by this account, we also found that there was significant variability *within* the different verb classes: Figure 5 shows the variable rates of EV2 for the 21 speech act predicates in our data set. Note that similar variation was found across the other verb classes as well.

We take this as evidence against this type of strong lexical licensing account, whereby membership of a given lexical class is what determines whether EV2 is available or not.

In Sections 3.1 and 3.2, we tested the predictions made by the two types of "selection-based accounts" discussed in Section 1.4 against large-scale data from the BRF corpus: one according to which V2 should not be available in the complements of factive verbs; and one whereby V2 is available, but entirely optional, in the complements of certain predicate types (1), but not others (2). We found that for neither of the two accounts were their predictions straightforwardly borne out. Rather, the distribution illustrated in Figure 4 suggests to us that, in addition to the lexical semantics of the embedding predicate, discourse factors play a significant role in driving the distribution of EV2, given that the different corpora can be understood to represent different discourse types. In particular, the distribution we observe looks like what we would expect if it were the case that EV2-clauses are associated with some kind of pragmatic meaning; the use of which is influenced by (but not solely determined by) the meaning of the embedding predicate, along with the type of discourse context in which the sentence is uttered. In the following section, we suggest that this pragmatic meaning is whether or not the embedded proposition p is discourse-new (Section 4). Subsequently we present further experimental (Section 5) and corpus (Section 6) results supporting this hypothesis.

4 EV2 and discourse novelty

To account for the interaction of discourse context and lexical semantics illustrated in Figure 4, we propose that:

(27)   a.   EV2-clauses have some interpretive effect. The distribution or use of this interpretive effect is influenced both by:
            i. the meaning of the embedding predicate;
            ii. the type of discourse context in which the sentence is uttered.

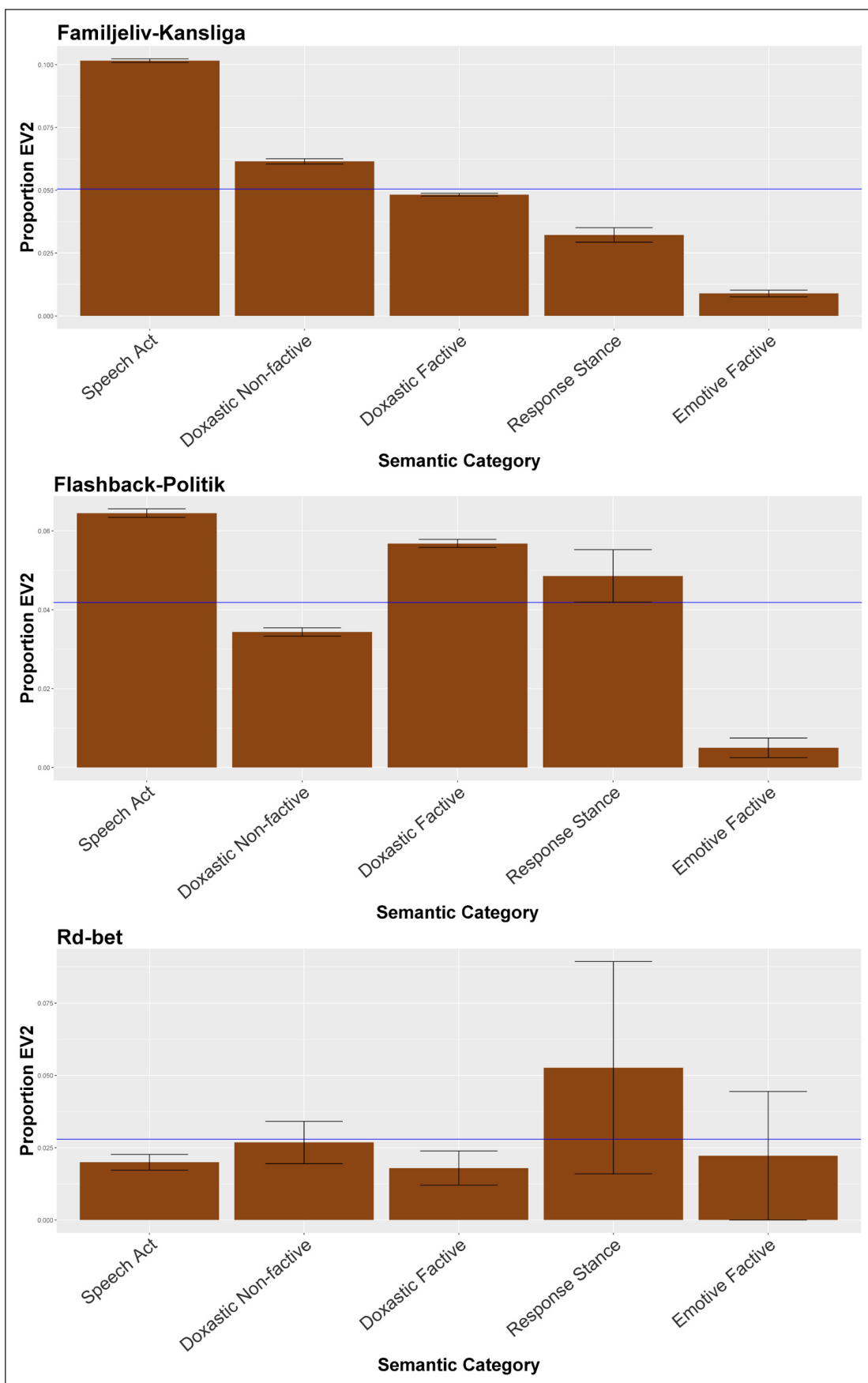   b.   The proposition denoted by an EV2 clause is interpreted as constituting discourse-new information.

Figure 4: Rates of EV2 from three of the BRF-corpora. From top to bottom: Familjeliv-känsliga (family-oriented discussion forum; 5,971,907 sentences), Flashback-Politik (online forum for political discussion; 2,841,872 sentences), and Rd-bet (government texts; 372,054 sentences).
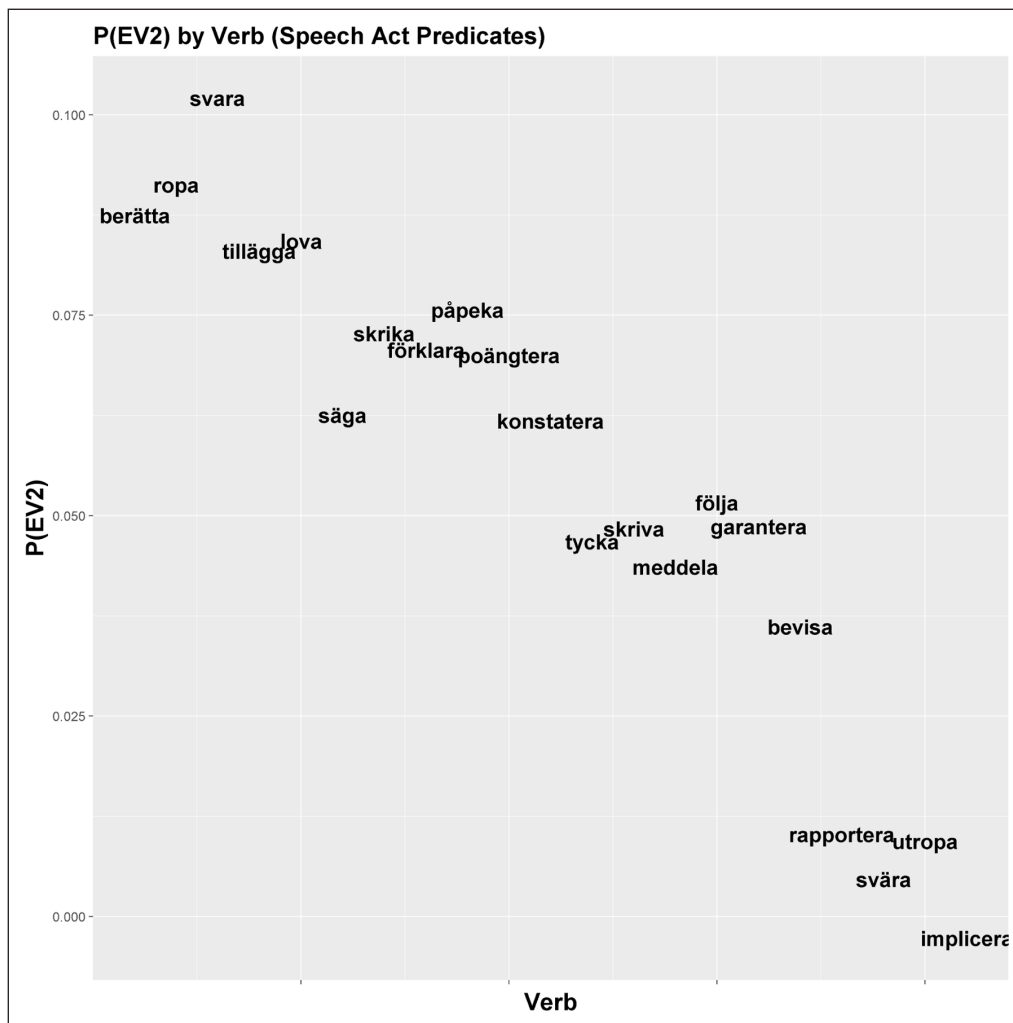
Figure 5: Probability of EV2 by lemma within the class of speech act verbs (the x-axis represents the 21 different verbs in this class ordered by proportion of EV2); plot based on data from a corpus of text from a political online forum (Flashback-Politik; 2,841,872 sentences).

Initial motivation for this proposal comes from considering the kinds of discourse contexts in which the relevant predicate types can felicitously be used. We observe that the different types of predicates vary in their ability to introduce entirely new information into the discourse; essentially, whether or not p has been previously discussed by the speaker and hearer. As shown in (28), this ability appears to correlate with the availability of EV2.[19]

(28)   [Uttered out of the blue:]
        Guess what — / You know what —
        a.   John **told me** that [$_p$ Bill and Anna broke up].   ✓V2
        b.   John **thinks** that [$_p$ Bill and Anna broke up].   ✓V2
        c.   John **discovered** that [$_p$ Bill and Anna broke up].   ✓V2
        d.   #John **appreciates** that [$_p$ Bill and Anna broke up].   ✗V2
        e.   #John **doubts** that [$_p$ Bill and Anna broke up].   ✗V2

Like the approach of Jensen & Christensen (2013), which we discussed and rejected in Section 1.4 on independent grounds, this proposal also relies on the notion of a Common

---

[19] Importantly, the # refers to the readings where p is presented to the hearer as discourse new information (as opposed to where the sentence makes a comment about the attitude holder). The same goes for (30) and (31).

Ground update. However, the type of context update is different in the two cases. In fact, the proposal advanced here is similar to that of Haegeman & Ürögdi (2010); Haegeman (2014); Kastner (2015) a.o., in that EV2 is taken not to be licensed in contexts where the embedded proposition p is discourse-old information.[20,21] However, the current proposal differs crucially in terms of our assumptions about factive predicates: whereas the above authors take all factive predicates to require p to be Common Ground, we follow Simons (2007) in her claim that the two components of factivity (that p is true, and that p is Common Ground) must be dissociated.[22] Simons' argumentation builds on question/answer sequences of the type we saw in (20). However, (28c) makes the same point: here, the embedded proposition p is clearly taken to be true by the speaker. However, there is no sense in which p is taken to be Common Ground. Note further that although our proposal relates crucially to the notion of a Common Ground update (it is at least plausible that speakers contribute new information to a discourse in an attempt to add that information to the Common Ground), it is not the case that EV2 is ruled out *only* in contexts in which p is Common Ground. The response predicates (like in (28e)) make this case the most clearly: here, p cannot be understood to be discourse-new information; however, there is also no sense in which the speaker is necessarily committed to p. What seems to be required here, for p to be discourse old in the sense that is relevant here, is that ?p (i.e., {p, ¬p}) is present as a question in the discourse. In that sense then, we take p to be discourse-old, though not Common Ground. In relation to this point, we might also note that the notion of discourse update that we have in mind is in fact different from that in Simons (2007) (illustrated in (20) and (23a)–(23b)): here, the embedded proposition provides a Common Ground update, relative to a particular Question Under Discussion. As we saw in our discussion in Section 1.4, this distinction does not seem to be what is relevant to the licensing of EV2. The pragmatic notion we propose to be relevant to EV2 is in fact a stronger notion of discourse novelty, where not just the proposition p itself is new to the discourse, but where the question of *whether p?* constitutes a discourse new *issue*.

---

[20] It is worth noting that the idea that discourse novelty is relevant to the licensing or availability of EV2 is implicit also in a number of "assertion-based" accounts (e.g. Julien 2009; Gärtner & Michaelis 2010; Woods 2016a; b). However, on these accounts, discourse novelty is typically regarded as a *secondary* condition, in addition to something like main point status, or belief(p). The main difference, then, between previous work and the current proposal, is that we don't appeal to any additional pragmatic factors beyond discourse novelty.

[21] Note that the notion of discourse novelty that is relevant here is different from that involved in cases of (*global*) accommodation (see for instance Karttunen 1974; Stalnaker 1974; Heim 1983; Thomason 1990; Van der Sandt 1992; Stalnaker 2002; Klinedinst 2010; Abrusán 2016). In the latter case, the speaker treats the relevant proposition, p, as discourse old (or presupposed) information, such that that the hearer will need to saliently adjust their Common Ground to include p, in order for the presupposition to be met by the context. In the case we have in mind, however, p is explicitly presented as discourse new information, and intended by the speaker to be interpreted as such (rather than, say, as a reminder; cf. discussion in Julien 2009). To this point, an anonymous reviewer points to the consequence of degree sentence in (1) below as a potential counterexample to our claim that Swedish EV2 is licensed by discourse novelty: the idea being that the proposition 'the fines do not tempt the proprietor to continue' could represent discourse new information, and yet, EV2 is not permitted here, as shown in (1b).

(1) *Norwegian* (Julien 2015: 161)
     a.   Bøtene   skal være så store at   de   ikke frister innehaveren   til å fortsette.
          fines.DEF shall be   so large that they not tempt proprietor.DEF to to continue
          'The fines should be so large that they do not tempt the proprietor to continue.'

     b. *Bøtene   skal være så store at   de   frister ikke innehaveren   til å fortsette.
          fines.DEF shall be   so large that they tempt not proprietor.DEF to to continue

However, this point speaks directly to the difference between on the one hand, global accommodation, and on the other, presenting a proposition p *as new information* in a move to update the context with p. We interpret (1) as involving the former. The question of how the discourse pragmatics interacts with different kinds of embedding environments more broadly, including different kinds of adverbial constructions and modal environments, is unfortunately beyond the scope of the current discussion.

[22] See also Djärv (2019a; b) for discussion and experimental evidence.

However, the lexical semantics of the embedding predicate is only *one* factor that constrains the ability of an embedded proposition to be presented as discourse-new information. The type of discourse, along with other properties of the sentence, matter too. The following example from the Flashback-Politik corpus (a political forum), involving the response verb *acceptera* ('accept') illustrates the latter point:

(29)   *Swedish*
    a.   Kan du inte bara slappna av och acceptera att   socialisterna **kan** inte
           can you not just chill   out and accept   that socialists.DEF can not
           vinna alla gånger?
           win   all   times
           'Why can't you just relax and accept that the socialists aren't going to win every time?'
    b.   Acceptera att   du kan inte älska alla   men du **kan** inte hata
           accept   that you can not love everyone but you can not hate
           alla   heller
           everyone either
           'Accept that you can't love everyone, but you can't hate everyone either.'

What appears to be happening in these cases is indeed that the speakers are presenting the embedded propositions ('the socialists can't win every time', and 'you can't love everyone, but you can't hate everyone either') as new information, in an attempt to update the Common Ground.

If the relevant dimension is truly the discourse status of the embedded proposition, the issue arises of how to test the hypothesis against corpus data, given that there is no direct way of measuring the discourse status of a given proposition in a corpus—especially not in one of this scale. However, it turns out that we *can* test whether or not the embedded proposition may constitute discourse-new information in a way that is quantifiable—but nevertheless independent of the identity of the matrix predicate—thus providing an independent test for our hypothesis. What we observe is that the speech act predicates and the doxastic non-factives, under negation, take on the property of requiring their complement to be discourse-old (similarly to the response predicates and the emotive factives), as illustrated in (30).

(30)   [Uttered out of the blue:]
    Guess what — / You know what —
    a.   John **told me** that [$_p$ Bill and Anna broke up].
    b.   John **thinks** that [$_p$ Bill and Anna broke up].
    c.   #John **didn't tell me** that [$_p$ Bill and Anna broke up].
    d.   #John **doesn't think** that [$_p$ Bill and Anna broke up].
    e.   #John **appreciates** that [$_p$ Bill and Anna broke up].
    f.   #John **doubts** that [$_p$ Bill and Anna broke up].
    g.   #John **doesn't appreciate** that [$_p$ Bill and Anna broke up].
    h.   #John **doesn't doubt** that [$_p$ Bill and Anna broke up].

Of course, as has been observed in previous work (e.g. Truckenbrodt 2006; Gärtner & Michaelis 2010), negating these verbs also negates their belief components. This has been taken to support a view such as that discussed in Section 1.4, whereby V2 is licensed by a belief context. However, we saw above that V2 disjunction presents a problem for any version of this view. The emotive factives provide an additional problem for that approach, given that in both positive and negative contexts, do they give rise to the obligatory inference that the matrix subject, and typically also the speaker, believe that the embedded

proposition is true (a property known as *projection*). Nevertheless, these verbs show the lowest rates of EV2 in the BRF-corpora (see Figure 4 in Section 3.2).

Based on this observation then, our hypothesis now predicts that the speech act and non-factive doxastic predicates, when negated, should show equally low rates of EV2 as the response predicates and the emotive factives (in both polarities), as shown in (31).

(31)   EV2: predicted distribution (verb type × negation interaction)
  a. John **told me** that [$_p$ Bill and Anna broke up].   ✓V2
  b. John **thinks** that [$_p$ Bill and Anna broke up].   ✓V2
  c. #John **didn't tell me** that [$_p$ Bill and Anna broke up].   ✗V2
  d. #John **doesn't think** that [$_p$ Bill and Anna broke up].   ✗V2
  e. #John **appreciates** that [$_p$ Bill and Anna broke up].   ✗V2
  f. #John **doubts** that [$_p$ Bill and Anna broke up].   ✗V2
  g. #John **doesn't appreciate** that [$_p$ Bill and Anna broke up]. ✗V2
  h. #John **doesn't doubt** that [$_p$ Bill and Anna broke up]. ✗V2

Before testing these predictions in the BRF-corpus, we wanted to make sure that this was indeed a robust property of these *predicate classes*, beyond our own intuitions about the particular verbs in (31). To this end, we carried out an experimental judgment task, which we describe in the following section.

5 Experiment: Negation & discourse novelty

The predictions illustrated in (31) are based on the observation that the speech act predicates and the doxastic non-factives, under negation, require their complement to be discourse-old. To make sure that this observation is empirically robust, we ran an experiment probing the effect of negation on whether or not p can be interpreted as discourse-new information under the different predicates types.

5.1 Methods

5.1.1 Design and materials

The experiment employed the "Guess what" test used above; here, framed in the context of a conversation between two friends, as shown in (32).

(32)   Two friends, Tom and Sue, run into each other. Tom says to Sue:
  Guess what! I just ran into Aaron, and he VERBS/DOESN'T VERB that
  [$_p$ Joel left his wife].

To measure the perceived discourse status of p, the participants were asked to complete a statement in which they had to rate on a Likert scale how likely they thought it was that the speaker and the hearer had talked about p before (7 = not likely; 1 = very likely), as shown in Figure 6.

Since our predictions were specifically about the interaction of negation with the speech act predicates and the doxastic non-factives, compared to the emotive factives and the response predicates, we did not include the doxastic factives in this experiment. We included three verbs from each lexical class:[23]

---

[23] The doxastic factives were excluded for the purpose of keeping the experiment size manageable; we had no specific prediction about how they should interact with negation. Note though, that recent work by Djärv (2019a; b) has replicated the results of the current study, including also the doxastic factives, in both German and Swedish: she found that the doxastic factives pattern with the speech act predicates for discourse novelty ratings (including the interaction with negation), and further, that these two verb classes (± negation) pattern alike also in terms of the acceptability of EV2.

> **Two friends, Tom and Sue, run into each other. Tom says to Sue:**
>
> *Guess what! I just ran into Aaron, and he denies that Joel left his wife.*
>
>
> It is **not likely** ○ ○ ○ ○ ○ ○ ○ **very likely**
>
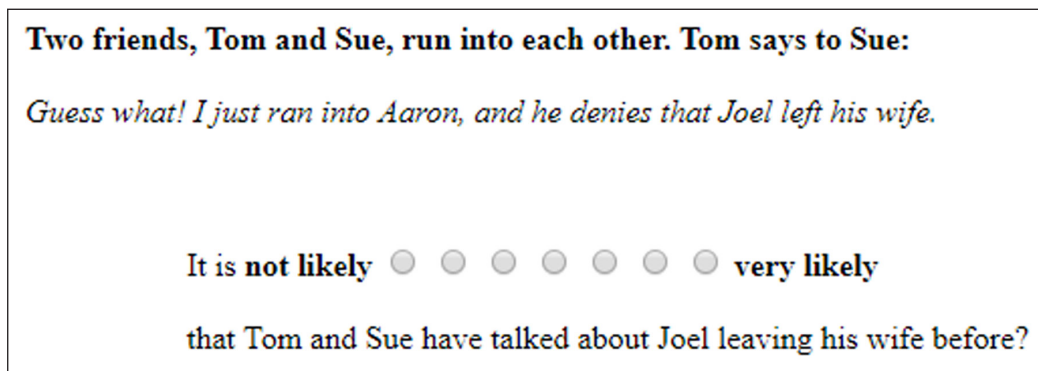> that Tom and Sue have talked about Joel leaving his wife before?

Figure 6: Screenshot of an experimental trial.

(33) a. Speech act predicates: *say, mention, tell (me)*
      b. Doxastic non-factives: *believe, think, assume*
      c. Response predicates: *accept, deny, admit*
      d. Emotive factives: *appreciate, regret, resent*

The experiment included 24 critical items, and 24 fillers, plus two practise items that were excluded from the analysis. Each item consisted of one verb and one (unique) complement clause, with variations in the two polarity conditions: positive (no matrix negation) vs. negative (with matrix negation). Whereas each embedded clause content occurred only in one item, every verb occurred in two items, so that each participant would see all conditions; [speech act vs. doxastic non-factive vs. emotive factive vs. response] × [negative vs. positive], across items, but with the specific content of the embedded clause shown only in one condition, counterbalanced across subjects using a latin-square design. Each subject thus saw each verb twice, once in the negative and once in the positive polarity (with different contents for the interlocutors and embedded clauses). Since there were three verbs per verb class, each participant saw each predicate type six times (three positive and three negative).

  We also included baseline floor and ceiling conditions for discourse-old vs. new status, as illustrated in (34). There were eight items of each kind.

(35) Control conditions:
      a. Discourse-new baseline (predict high ratings):
          Guess what! Joel left his wife.
      b. Discourse-old baseline (predict low ratings):
          Guess what! John thinks, like you do, that Joel left his wife.

Additionally, the experiment included eight pure fillers, involving conditionals (35). For these, the participants rated the likelihood of the proposition in the antecedent being old vs. new (here, that Nadine travelled to Asia).

(36) *Guess what! I just ran into Lisa, and she said that if Nadine travelled to Asia, then she must have lots of interesting stories to tell.*

Importantly, the *Guess what* experiment was run in English rather than Swedish with translations of the original predicates. While translations are inherently noisy—fine-grained denotations or associations may differ cross-linguistically—what's crucial is that status in particular predicate classes of interest is held constant. Conducting the experiment in English provides a more rigorous test of the formal properties under consideration. We remove any potential lexically specific confounds present between acceptability judgements in English

and rates of EV2 in Swedish; the only properties shared between translations are abstract semantic ones rather than Swedish specific distributional information (frequency, rates of EV2, etc.) Any additional noise imparted through the translation process can only make it more difficult to establish statistical significance to this connection rather than easier (hence providing a test which is both theoretically and empirically stronger).[24]

The experiment was implemented in *Ibex,* and took 10–15 minutes to complete. An archived version of the experiment is available on: http://spellout.net/ibexexps/ SchwarzLab/DiscFam.Archive/experiment.html?id=archive.

5.1.2 Participants

56 undergraduate students, recruited through the University of Pennsylvania's Psychology Department's subject pool (SONA), participated in the study for course credit. They were given a link to the experiment to take it online in their own time. Based on responses in the control conditions, we excluded the responses from five participants who appeared to have reversed the scale, leaving us with the responses from 51 participants.

5.1.3 Analysis

The data was analyzed in R (version 3.5.0). To test our predictions, we carried out a regression by fitting a linear mixed effects model, using *lmer* from the *lme4* package. The package *lmerTest* was used to generate p-values. The dependent variable was the perceived likelihood of p being new information. The model included Predicate Type, Polarity type, and their interaction (base levels: predicate type = Speech Act; polarity = Positive) as fixed effects. It also included a random intercept for participant and item. We also ran a model predicting the responses from the individual predicates (Verb Lemma). The conditional fillers (35) were excluded from the analysis.

To identify outliers we created two sets of subjects based on their responses in the two control conditions (34): (a) subjects whose average response were more than one standard deviation below the mean in the discourse-new condition, and (b) subjects whose average response were more than one standard deviation above the mean in the discourse-old condition. We then took the intersection of the two sets, thus giving us only the participants who were outliers for *both* control conditions (n = 5). Thus, the subjects that we excluded from the analysis were those who deviated from the mean by more than one standard deviation in the "unexpected" directions for the two control conditions. To compare the data with and without the outliers, we used *r.squaredGLMM* from the *MuMin* package, to calculate the (marginal and conditional) R squared values for a model with the full data set (n = 56), and the subsetted data set (n = 51), to determine how well the model fits the data. R squared is a statistical measure of how close the data are to the fitted regression line (R squared = Explained variation/Total variation). The data was plotted using *ggplot* from the *ggplot2* package; error bars represent the standard error of the mean.

5.1.4 Predictions

We predict that matrix negation will interact with predicate type, such that the speech act predicates and the doxastic non-factives receive significantly higher ratings in the positive than in the negated condition. We predict that the response predicates and the emotive factives should receive low ratings in both polarity conditions.

---

[24] As an anonymous reviewer points out, it has been noted, for instance by Wiklund et al. (2009), that the counterpart of the English verb meaning roughly 'regret' differ in certain ways the across the Scandinavian languages; particularly that Icelandic *harma* differs in certain ways from Swedish *ångra* and Norwegian *angre.* The important point here is that the Swedish and the English counterparts nevertheless share the relevant pragmatic property of requiring the embedded proposition p to be Given; an empirical assumption that is supported by resent experimental results by Djärv (2019a; b).

5.2 Results

Figures 7–8 show the responses for the critical items and the two control conditions (34) (the responses for the conditional fillers (35) are not included).

The R squared values for the data with and without the outliers are given in Table 5. As expected, we find that with the subsetted data (without the outliers) the model fits the data better than with the full data set. This is true both for the models based on predicate type and verb lemma. We also observe that for none of the models is there a big difference between the conditional and the marginal R squared values, showing us that most of the variation in the data is explained by the fixed effects.

The linear mixed effects model (based on predicate type, without outliers, n = 51) shows a main effect of predicate type. Relative to the intercept (6.0920; this is the mean
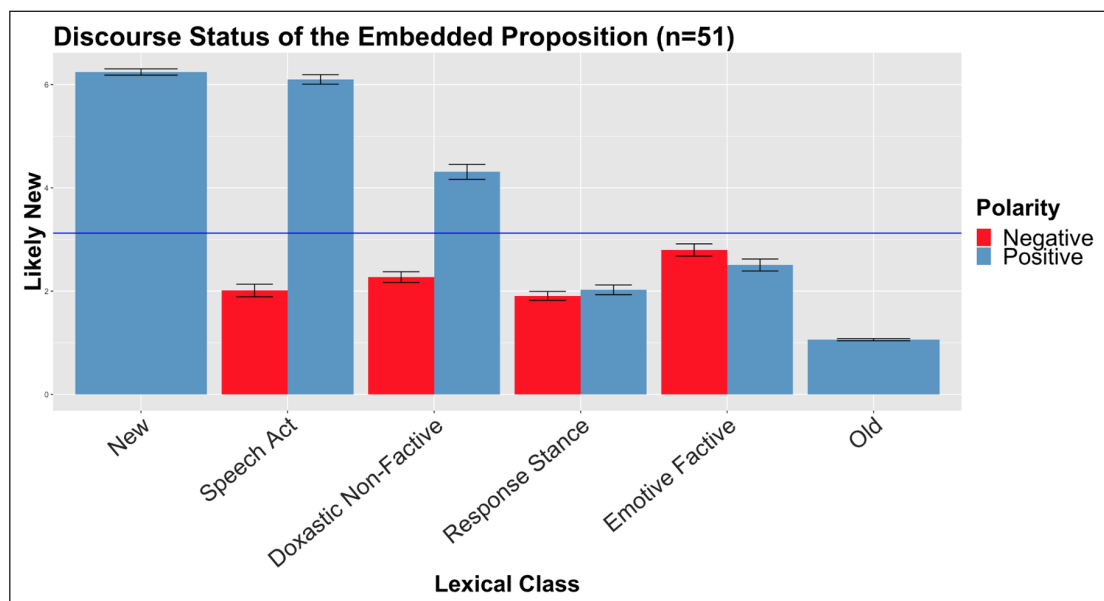
Figure 7: Response patterns by predicate type and polarity (critical and control conditions). The blue horizontal line shows the overall mean.
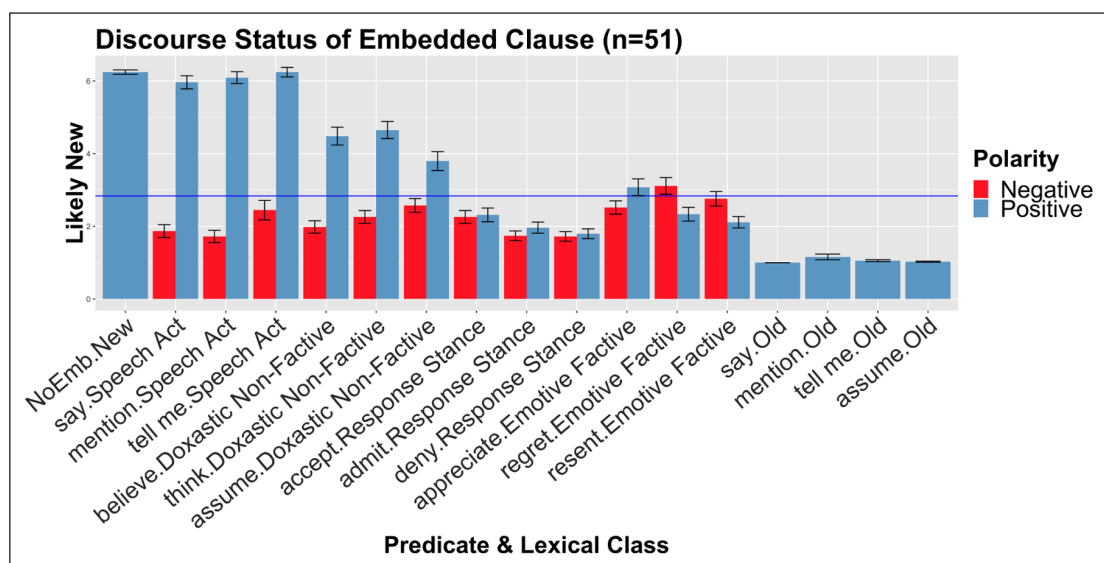
Figure 8: Response patterns by predicate and polarity (critical and control conditions). The blue horizontal line shows the overall mean

Table 5: R squared values for the data set with and without outliers. Marginal R squared values consider only the fixed effects; the conditional R squared values consider both the fixed and the random effects.

Data	Marginal R²	Conditional R²
Predicate Type: With outliers (n = 56)	0.67	0.71
Predicate Type: Without outliers (n = 51)	0.71	0.75
Verb Lemma: With outliers (n = 56)	0.61	0.65
Verb Lemma: Without outliers (n = 51)	0.72	0.75

of the dependent variable for the two base levels: predicate type = Speech Act and polarity = positive), the model shows that the following conditions are significantly different ($p < 0.001$) (the numbers represent the model estimated difference relative to the base levels): Doxastic Non-factive ($\beta = -1.7744$); Response ($\beta = -4.1227$), Emotive Factive ($\beta = -3.5806$), old information controls ($\beta = -5.0297$), and polarity ($\beta = -4.0726$, $p < 0.001$). The new information controls did not differ significantly from the base levels ($\beta = 0.1532$, $p = 0.361$).

The model shows the following significant interactions ($p < 0.001$): the difference between the positive and the negative polarity is greater for the Speech Act predicates than for the other predicate types; Doxastic Non-factives ($\beta = 2.0267$), Response ($\beta = 3.9924$), and Emotive Factives ($\beta = 4.3679$). Given the fixed effect of polarity we just observed ($\beta = -4.0726$, $p < 0.001$), this means that the difference between the two polarity conditions for the Doxastic Non-factives is about half the size of that for the Speech Act predicates, whereas for the Response predicates and Emotive factives, there is essentially no difference between the two polarities: in these conditions, the effect of negation is close to zero. In fact, the Emotive Factives appear to show a small difference in the opposite direction from the other conditions.

These results then are precisely what we predicted (Section 5.1.4). Additionally, the difference between the Speech Act and Non-factive Doxastic predicates is in line with the observation that the Speech Act predicates show the overall highest levels of EV2. By testing acceptability judgements via English translations rather than the original Swedish we have ensured that any lexically-specific behavior of individual English predictes is precisely limited to English. Since the only connection between the English test items and their Swedish counterparts is through their formal semantic properties, we can rest assured that the strong connection we see between discourse-novelty (tested in English) and rates of EV2 (evaluated in Swedish) is robust. We know from this that the connection is due to structural causality rather than something psycholinguistic in nature like learned co-variation.

6 Testing our prediction: EV2 negation effect

Having confirmed that matrix negation independently impacts the interpretation of the embedded proposition as discourse-old vs. -new information, for the Speech Act predicates and the Doxastic Non-factives, we were able to test our prediction that the rates of EV2 in the corpus should be notably lower for the negated Speech Act and Doxastic Non-factive predicates, than for their non-negated counterparts. As shown in Figure 9, this prediction was borne out. This effect was confirmed by a Wilcoxon Rank Sum test ($W = 749$, $p < 0.008$), and holds across all corpora we looked at.

Importantly, this was not due to a main effect of negation, but reflects specifically the interaction of negation and the speech act and non-factive doxastic predicates, as predicted from the experimental results in Section 5. We also predicted that negation should not significantly impact the rates of EV2 for the Response stance predicates and the Emotive

factives, as is borne out in Figure 10. This (lack of) effect was confirmed by a Wilcoxon Rank Sum test (W = 133, p = 0.7322).

It is also worth pointing out the one place in the current data where our proposal makes clearly different predictions from the type of account discussed in Section 1.4, that takes EV2 to be licensed by the presence of a belief that p (e.g. Truckenbrodt 2006). On this view, we should expect to see an asymmetry between the positive and the negative response stance verbs (e.g. *accept/admit* vs. *doubt/deny*), as well as an interaction with negation. In
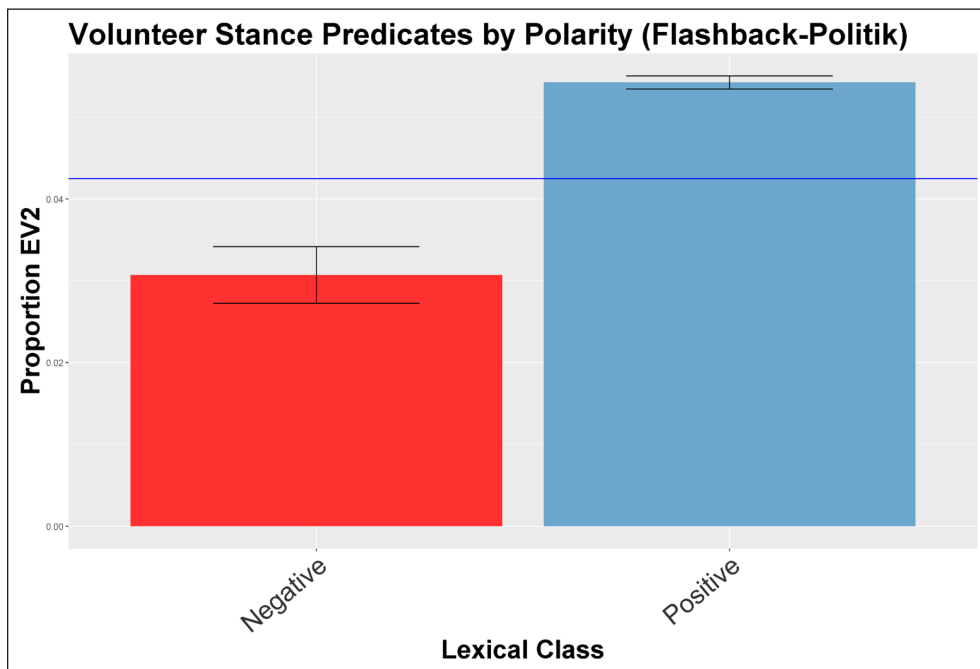
Figure 9: Rates of EV2 with the speech act predicates and the doxastic non-factives under negative and positive polarity.
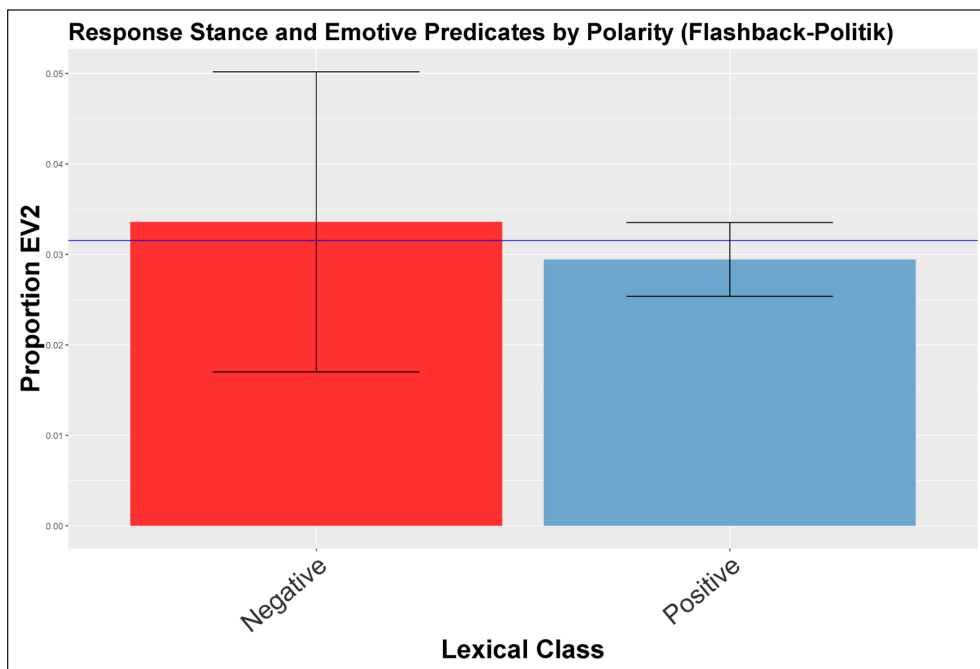
Figure 10: Rates of EV2 with the response stance predicates and emotive factives under negative and positive polarity.

particular, this hypothesis predicts: (i) that EV2 should be possible under the positive, but not the negative response predicates; and (ii) that the *negated* positive response predicates should show lower rates of EV2 than the non-negated ones, and vice versa for the negative response predicates. As shown in Figure 8 however, discourse novelty does not vary drastically along the positive/negative dimension. We therefore do not predict that these should vary with respect to the availability of EV2. Looking at the rates of EV2 under the response predicates in the BRF corpus, we observe no clear difference between the positive and the negative response predicates. A Wilcoxon Rank Sum test shows no significant difference in EV2 under negated vs. positive response stance verbs (W = 13, p = 0.4396). This supports the view proposed in the current paper, whereby EV2 is licensed in contexts where p is treated as discourse new information.

7 Conclusions

The findings presented in this paper support our hypothesis, outlined in Section 4, that EV2 is licensed in contexts where the embedded proposition constitutes discourse-new information. Importantly, this is a pragmatic property of an utterance in context—constrained, but not determined, by the lexical semantics of the matrix predicate. Other factors that play a role include the pragmatic context of the utterance, as well as other grammatical properties of the sentence. Here we investigated the effect of one such factor, namely matrix negation, and showed that certain predicates interact with negation in a way that constrains the potential discourse-status of a sentence. These results then made novel predictions regarding the distribution of embedded verb second in the corpus, which we showed were borne out. Note that while we only looked at the interaction with negation, the naturally occurring sentences in (29), from the BRF corpus, suggest that negation is only one potentially relevant grammatical factor.

In addition to the effect of discourse novelty, we also observed that the rates of EV2 are graded by formality, such that rates of EV2 are much lower in written, formal contexts. This replicates results from Heycock & Wallenberg (2013), and is in line with the observation that (at least in Swedish) there exists a prescriptive bias against EV2.[25] It remains a question for future work to tease out in more detail to what extent these types of factors are responsible for the overall low rates of EV2 in our corpus data. A further issue that remains for future work to address is whether the observation that EV2 in Swedish is licensed by discourse novelty can be extended to EV2 in other syntactic frames (for instance in cases where the pre-verbal element is a focused or topicalized non-subject XP, or in different kinds of adverbial clauses). We also noted that our account of Swedish EV2 appears to be similar to that proposed by Gärtner & Michaelis (2010) in their analysis of German main clause V2. This then provides some support for a homogeneous account, not only for embedded and matrix clause V2 (contrary to e.g. Truckenbrodt 2006), but also for V2 across languages (see also Djärv 2019a; b for experimental results supporting this position). A final issue for future investigation concerns the question of to what extend the current account generalizes to other MCP. For recent theoretical and experimental work on variation among different types of MCP, see Jacobs (2018) and Djärv (2019a; b).

It's worth noting that while previous work has pointed to both discourse familiarity and negation as factors relevant to the licensing of EV2 and MCP, their effects have been

---

[25] As an anonymous reviewer points out, the prescription against EV2 is quite explicit in the Swedish educational system at all levels, and is commonly referred to as the "BIFF"-rule. See for instance https://sfipatxi.wordpress.com/2018/01/05/ordfoljd-i-bisats/#more-892.

interpreted disjunctively, as evidence for different theoretical accounts. On an account where EV2 is licensed by the presence of a belief-context, the effect of negation on EV2 is taken to follow from the negation of the attitude holder's belief that p. However, we saw above that this view over-generates. On the kind of lexical licensing approach advocated by Haegeman and colleagues, MCP are claimed to be blocked by "presuppositionality" or "referentiality" (in their terminology). However, they take this to involve all factive verbs, and make no reference to negation. Here, we link the interaction of verb-type and negation explicitly to the discourse status of p, thus getting a unified account of the effect of negation and the role of predicate type. However, neither (speaker or attitude holder) belief, nor factivity plays any explanatory role on our proposal.

Apart from contributing to the empirical picture and the theoretical debate regarding EV2 and Main Clause Phenomena more broadly, the present study also represents a methodological contribution. Syntactic "optionality" is not an inherently unified phenomenon. While (particularly lexical-level) usage rates could potentially result from probabilistic representations, this need not be the case as specific output statistics can emerge from an interaction with context. We need to be careful in our interpretation of usage data and evaluate multiple theories (including both grammatical and psycholinguistic ones) when applicable. Yet, despite these caveats, we are still able to learn a good deal about grammatical representation directly from observational usage statistics.

Abbreviations

BFR = Språkbanken Corpus of Swedish (Borin, Forsberg and Roxendal), CG = Common Ground, DEF = definite suffix, EC = embedded clause, EV2 = Embedded Verb Second, MC = main clause, MCP = Main Clause Phenomena, P = proposition.

Additional Files

The additional files for this article can be found as follows:

- **EV2 Corpus Rates.** Corpus usage data by individual predicate; including EV2, frequency, polarity-dependent behavior, etc. DOI: https://doi.org/10.5334/gjgl.867.s1
- **Guess What Data.** Raw output data from the "Guess What" discourse status experiment. DOI: https://doi.org/10.5334/gjgl.867.s2

Acknowledgements

Funding Information

Competing Interests

The authors have no competing interests to declare.

Author Contributions

Spencer Caplan and Kajsa Djärv authors contributed equally to this manuscript.

References

Abrusán, Márta. 2011. Presuppositional and negative islands: A semantic account. *Natural Language Semantics* 19(3). 257–321. DOI: https://doi.org/10.1007/s11050-010-9064-4

Abrusán, Márta. 2016. Presupposition cancellation: Explaining the 'soft—hard' trigger distinction. *Natural Language Semantics* 24(2). 165–202. DOI: https://doi.org/10.1007/s11050-016-9122-7

Adams, Marianne. 1985. Government of empty subjects in factive clausal complements. *Linguistic Inquiry* 16(2). 305–313.

Aelbrecht, Lobke, Liliane Haegeman & Rachel Nye. 2012. *Main Clause Phenomena: New Horizons* (Linguistik Aktuell/Linguistics Today), vol. 190. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.190

Andersson, Lars-Gunnar. 1975. *Form and function of subordinate clauses*. Göteborg: University of Göteborg dissertation.

Bentzen, Kristine. 2010. Exploring embedded main clause phenomena: The irrelevance of factivity and some challenges from V2 languages. *Theoretical Linguistics* 36(2/3). 163–172. DOI: https://doi.org/10.1515/thli.2010.010

Bentzen, Kristine, Gunnar Hrafn Hrafnbjargarson, Þorbjörg Hróarsdóttir & Anna-Lena Wiklund. 2007. The Tromsø guide to the Force behind V2. *Working Papers in Scandinavian Syntax* 79. 93–118.

Bhatt, Rajesh. 2010. Comments on "Referential CPs and DPs: An operator movement account". *Theoretical Linguistics* 36. 173–177. DOI: https://doi.org/10.1515/thli.2010.011

Bianchi, Valentina & Mara Frascarelli. 2009. Is topic a root phenomenon? Published on LingBuzz; reference lingbuzz/000954.

Bock, Kathryn & Willem Levelt. 2002. Language production: Grammatical encoding. In Gerry T.M. Altmann (ed.), *Psycholinguistics: Critical concepts in psychology* 5. 405–464. New York: Routledge.

Bogal-Allbritten, Elizabeth & Keir Moulton. 2018. Nominalized clauses and reference to propositional content. In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of Sinn und Bedeutung 21* 21. 215–232.

Borin, Lars, Markus Forsberg & Johan Roxendal. 2012. Korp – the corpus infrastructure of Språkbanken. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Mehmet Ugur Dogan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, 474–478. Istanbul: European Language Resources Association (ELRA).

Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in search of its evidential base* 96. 77–96. Walter de Gruyter.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation (Proceedings of the Amsterdam Colloquium)*, 69–94. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen.

Brill, Eric. 2000. Part-of-speech tagging. In Robert Dale, Hermann Moisl & Harold Somers (eds.), *Handbook of natural language processing*, 403–414. New York: Marcel Dekker.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language*, 711–733. DOI: https://doi.org/10.1353/lan.2006.0186

Bybee, Joan & David Eddington. 2006. A usage-based approach to Spanish verbs of 'becoming'. *Language* 82(2). 323–355. DOI: https://doi.org/10.1353/lan.2006.0081

Caplan, Spencer. 2018. Incremental generation drives "efficient" language production. *Paper presented at AMLaP 2018 Architectures and Mechanisms for Language Processing*. Berlin, Germany.

Dayal, Veneeta & Jane Grimshaw. 2009. Subordination at the interface: The Quasi-Subordination hypothesis. Unpublished manuscript.

De Cuba, Carlos. 2017. In a referential manner of speaking. *Paper presented at the 41st Penn Linguistics Conference PLC41*. Philadelphia, PA.

De Cuba, Carlos & Barbara Ürögdi. 2009. Eliminating factivity from syntax: Sentential complements in Hungarian. In Marcel den Dikken & Robert M. Vago (eds.), *Approaches to Hungarian* 11. 29–63. Amsterdam/New York: John Benjamins. DOI: https://doi.org/10.1075/atoh.11.03cub

De Cuba, Carlos & Barbara Ürögdi. 2010. Clearing up the 'Facts' on Complementation. In Jon Scott Stevens (ed.), *Proceedings of the 33rd Annual Penn Linguistics Colloquium* (University of Pennsylvania Working Papers in Linguistics (PWPL)), 16. 41–50. Philadelphia, PA: ScholarlyCommons.

De Haan, Germen J. 2001. More is going on upstairs than downstairs: Embedded root phenomena in West Frisian. *The Journal of Comparative Germanic Linguistics* 4(1). 3–38. DOI: https://doi.org/10.1023/A:1012224020604

Den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In Werner Abraham (ed.), *On the formal syntax of the Westgermania*, 47–131. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.3.03bes

Djärv, Kajsa. 2019a. *Factive and assertive attitude reports*. Philadelphia, PA: University of Pennsylvania dissertation.

Djärv, Kajsa. 2019b. Propositional attitude reports: The syntax of presupposition & assertion. *Paper presented at Semantics and Linguistic Theory 29 (SALT29)* at University of California, Los Angeles, CA.

Djärv, Kajsa, Caroline Heycock & Hannah Rohde. 2017. Assertion and factivity: Towards explaining restrictions on Embedded V2 in Scandinavian. In Laura Bailey & Michelle Sheehan (eds.), *Order and structure in syntax (open generative syntax)*, 1–31. Berlin: Language Science Press.

Du Bois, John W. 1985. Competing motivations. In John Haiman (ed.), *Iconicity in syntax* 6. 343–365. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/tsl.6.17dub

Elliott, Patrick D. 2016. Explaining DPs vs. CPs without syntax. *Proceedings from the 52nd Annual Meeting of the Chicago Linguistic Society (CLS52)* 52(1). 171–185.

Elsness, Johan. 1984. That or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies* 65(6). 519–533. DOI: https://doi.org/10.1080/00138388408598357

Emonds, Joseph. 1970. *Root and structure-preserving transformations*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Ferreira, Victor & Gary Dell. 2000. Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive psychology* 40(4). 296–340. DOI: https://doi.org/10.1006/cogp.1999.0730

Gärtner, Hans-Martin. 2000. Are there V2 relative clauses in German? *The Journal of Comparative Germanic Linguistics* 3(2). 97–141. DOI: https://doi.org/10.1023/A:1011432819119

Gärtner, Hans-Martin. 2002. On the force of V2 declaratives. *Theoretical Linguistics* 28(1). 33–42. DOI: https://doi.org/10.1515/thli.2002.28.1.33

Gärtner, Hans-Martin & Jens Michaelis. 2010. On modeling the distribution of declarative V2-clauses: The case of disjunction. In Sebastian Bab & Klaus Robering (eds.), *Judgements and propositions*, 11–25. Berlin: Logos Verlag.

Gleitman, Lila, David January, Rebecca Nappa & John C. Trueswell. 2007. On the give and take between event apprehension and utterance formulation. *Journal of memory and language* 57(4). 544–569. DOI: https://doi.org/10.1016/j.jml.2007.01.007

Green, Georgia M. 1976. Main clause phenomena in subordinate clauses. *Language*, 382–397. DOI: https://doi.org/10.2307/412566

Grice, H. Paul. 1957. Meaning. *The philosophical review* 66(3). 377–388. DOI: https://doi.org/10.2307/2182440

Haegeman, Liliane. 2010. The movement derivation of conditional clauses. *Linguistic Inquiry* 41(4). 595–621. DOI: https://doi.org/10.1162/LING_a_00014

Haegeman, Liliane. 2012. The syntax of MCP: Deriving the truncation account. In Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds.), *Main clause phenomena: New horizons* 190. 113–134. Oxford: John Benjamins. DOI: https://doi.org/10.1075/la.190.05hae

Haegeman, Liliane. 2014. Locality and the distribution of main clause phenomena. In Enoch Oladé Aboh, Maria Teresa Guasti & Ian Roberts (eds.), *Locality* (Oxford Studies in Comparative Syntax), chap. 8. 186–222. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199945269.003.0008

Haegeman, Liliane & Barbara Ürögdi. 2010. Referential CPs and DPs: An operator movement account. *Theoretical Linguistics* 36. 111–152. DOI: https://doi.org/10.1515/thli.2010.008

Hale, John T. 2014. *Automaton theories of human sentence comprehension* (CSLI Studies in Computational Linguistics 14). Stanford, CA: CSLI Publications. DOI: https://doi.org/10.1353/lan.2016.0088

Hegarty, Michael. 1992. Adjunct extraction without traces. In Dawn Bates (ed.), *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 209–222. Stanford, CA: CSLI Publications.

Heim, Irene. 1982. *The semantics of definite and indefinite NPs*. Amherst, MA: University of Massachusetts at Amherst dissertation.

Heim, Irene. 1983. On the projection problem for presuppositions. In Daniel P. Flickinger (ed.), *Proceedings of WCCFL 2*, 114–125. Stanford, CA: CSLI Publications. DOI: https://doi.org/10.1002/9780470758335.ch10

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9. 183–211. DOI: https://doi.org/10.1093/jos/9.3.183

Heycock, Caroline & Joel Wallenberg. 2013. How variational acquisition drives syntactic change: The loss of verb movement in Scandinavian. *Journal of Comparative Germanic Linguistics* 16. 127–157. DOI: https://doi.org/10.1007/s10828-013-9056-0

Holmberg, Anders. 2015. Verb second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax–theory and analysis. An international handbook of contemporary syntactic research*, 342–383. Walter de Gruyter 2nd edn. DOI: https://doi.org/10.1515/9783110377408.342

Holmberg, Anders & Christer Platzack. 1991. On the role of inflection in Scandinavian syntax. In Werner Abraham, Wim Kosmeijer & Eric Reuland (eds.), *Issues in Germanic syntax*, 93–118. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110847277.93

Holmberg, Anders & Christer Platzack. 1995. *The role of inflection in the syntax of the Scandinavian languages*. Oxford: Oxford University Press.

Hooper, Joan. 1975. On assertive predicates. In John P. Kimball (ed.), *Syntax and semantics*, 91–124. New York: Academy Press.

Hooper, Joan & Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.

Hopper, Paul. 1987. Emergent grammar. In *Annual meeting of the Berkeley Linguistics Society* 13. 139–157. DOI: https://doi.org/10.4324/9780203809068.ch21

Iatridou, Sabine & Anthony Kroch. 1992. The licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. In *Working Papers in Scandinavian Syntax* 50. 1–24. Lund: Lund University.

Jacobs, Joachim. 2018. On main clause phenomena in German. In Markus Steinbach & Arnim Grewendorf, Güntherad von Stechow (eds.), *Linguistische Berichte heft* 254. 131–182. Helmut Buske Verlag.

Jaeger, T. Florian. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive psychology* 61(1). 23–62. DOI: https://doi.org/10.1016/j.cogpsych.2010.02.002

Jensen, Torben Juel & Tanya Karoli Christensen. 2013. Promoting the demoted: The distribution and semantics of "main clause word order" in spoken Danish complement clauses. *Lingua* 137. 38–58. DOI: https://doi.org/10.1016/j.lingua.2013.08.005

Jeoung, Helen. 2018. *Optional elements in Indonesian morphosyntax*. Philadelphia, PA: University of Pennsylvania dissertation.

Julien, Marit. 2009. Embedded clauses with main clause word order in Mainland Scandinavian. Published on LingBuzz; reference lingBuzz/000475.

Julien, Marit. 2015. The force of V2 revisited. *The Journal of Comparative Germanic Linguistics* 18(2). 139–181. DOI: https://doi.org/10.1007/s10828-015-9073-2

Karttunen, Lauri. 1971. Some observations on factivity. *Papers in Linguistics* 4. 55–69. DOI: https://doi.org/10.1080/08351817109370248

Karttunen, Lauri. 1974. Presuppositions and Linguistic Context. *Theoretical Linguistics* 1. 181–194. DOI: https://doi.org/10.1515/thli.1974.1.1-3.181

Kastner, Itamar. 2015. Factivity mirrors interpretation: The selectional requirements of presuppositional verbs. *Lingua* 164. 156–188. DOI: https://doi.org/10.1016/j.lingua.2015.06.004

Keenan, Edward L. 1971. Two kinds of presupposition in natural language. In Charles J. Fillmore & D. Terence Langéndoen (eds.), *Studies in linguistic semantics*, 45–54. New York, NY: Holt.

Kiparsky, Paul & Carol Kiparsky. 1970. Fact. In Michael Bierwisch & Karl Erich Heidolph (eds.), *Progress in Linguistics*, 143–173. The Hague: Mouton. DOI: https://doi.org/10.1515/9783111350219.143

Klinedinst, Nathan. 2010. Totally hardcore semantic presuppositions. Unpublished manuscript.

Maki, Hideki, Lizanne Kaiser & Masao Ochi. 1999. Embedded topicalization in english and japanese. *Lingua* 109. 1–14. DOI: https://doi.org/10.1016/S0024-3841(98)00055-2

McClosky, David, Eugene Charniak & Mark Johnson. 2010. Automatic domain adaptation for parsing. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 28–36. Association for Computational Linguistics.

Miyagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/10958.001.0001

Pesetsky, David. 1991. Zero syntax ii: Infinitival complementation. Unpublished manuscript.

Piantadosi, Steven T. 2014. Zipf's word frequency law in natural language: A critical review and future directions. *Psychonomic bulletin & review* 21(5). 1112–1130. DOI: https://doi.org/10.3758/s13423-014-0585-6

Platzack, Christer. 1987. The Scandinavian languages and the null subject parameter. *Natural Language and Linguistic Theory* 5. 377–401. DOI: https://doi.org/10.1007/BF00134554

Platzack, Christer & Anders Holmberg. 1989. The role of Agr and finiteness in Germanic VO languages. In *Working Papers in Scandinavian Syntax* 43. 51–76. Lund: Lund University.

Reis, Marga. 1997. Zum syntaktischen Status unselbstndiger verbzweit-stze. In Christa Drüscheid, Karl Heinz Ramers & Monika Schwarz (eds.), *Sprache im Fokus*, 112–144. Tübingen: Niemeyer.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar* (Kluwer International Handbooks of Linguistics), 281–337. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-011-5420-8_7

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In Jae-Hak Toon & Andreas Kathol (eds.), *Working Papers in Linguistics-Ohio State University Department of Linguistics* 49. 91–136. Columbus, OH: The Ohio State University.

Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5. 1–69. DOI: https://doi.org/10.3765/sp.5.6

Roland, Douglas, Jeffrey Elman & Victor Ferreira. 2006. Why is that? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition* 98(3). 245–272. DOI: https://doi.org/10.1016/j.cognition.2004.11.008

Rooryck, Johan. 1992. Negative and factive islands revisited. *Journal of Linguistics* 28. 343–374. DOI: https://doi.org/10.1017/S0022226700015255

Sekine, Satoshi. 1997. The domain dependence of parsing. In *Proceedings of the Fifth Conference on Applied Natural Language Processing (ANLC '97)*, 96–102. Washington, DC: Association for Computational Linguistics. DOI: https://doi.org/10.3115/974557.974572

Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 117. 1034–1056. DOI: https://doi.org/10.1016/j.lingua.2006.05.006

Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton Munitz & Peter Unger (eds.), *Semantics and philosophy*, 197–213. New York: New York University Press.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics* 9. 315–322. Cambridge, MA: Academic Press.

Stalnaker, Robert C. 2002. Assertion. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings*, 147–161. Wiley Online Library. DOI: https://doi.org/10.1002/9780470758335.ch5

Tamminga, Meredith, Laurel MacKenzie & David Embick. 2016. The dynamics of variation in individuals. *Linguistic Variation* 16(2). 300–336. DOI: https://doi.org/10.1075/bct.97.06tam

Thomason, Richmond H. 1990. Accommodation, meaning, and implicature: Interdisciplinary foundations for pragmatics. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 325–363. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/3839.001.0001

Truckenbrodt, Hubert. 2006. On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics* 32(3). 257–306. DOI: https://doi.org/10.1515/TL.2006.018

Van der Sandt, Rob A. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9(4). 333–377. DOI: https://doi.org/10.1093/jos/9.4.333

Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

Wechsler, Stephen. 1991. Verb second and illocutionary force. In Katherine Leffel & Dennis Bouchard (eds.), *Views on phrase structure*, 177–191. Dordrecht: Kluwer. DOI: https://doi.org/10.1007/978-94-011-3196-4_10

Weerman, Fred P. & Germen J. de Haan. 1986. Finiteness and verb fronting in Frisian. In Hubert Haider & Martin Prinzhorn (eds.), *Verb second phenomena in Germanic languages* (Publications in Language Sciences 21), 77–110. Dordrecht: Foris.

Wiklund, Anna-Lena. 2010. In search of the force of dependent verb second. *Nordic Journal of Lingusitics* 33(1). 81–91. DOI: https://doi.org/10.1017/S0332586510000041

Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson & Þorbjörg Hróarsdóttir. 2009. On the distribution and illocution of V2 in Scandinavian that-clauses. *Lingua* 119(12). 1914–1938. DOI: https://doi.org/10.1016/j.lingua.2009.03.006

Woods, Rebecca. 2016a. Embedded inverted questions as embedded illocutionary acts. In Kyeong-min Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel-Thompson & Lisa Shorten (eds.), *Proceedings of WCCFL 33*, 417–426. Sommerville, MA: Cascadilla Proceedings Project.

Woods, Rebecca. 2016b. *Investigating the syntax of speech acts: Embedding illocutionary force*. York: University of York dissertation.

Yang, Charles. 2013. Who's afraid of George Kingsley Zipf? or: Do children and chimps have language? *Significance* 10(6). 29–34. DOI: https://doi.org/10.1111/j.1740-9713.2013.00708.x