

INFORMATION-STRUCTURAL CONSTRAINTS ON WORD ORDER IN ADVANCED L2 SWEDISH

UTE BOHNACKER
(Uppsala University)

ABSTRACT: In a recent study of the prefield, i.e. the clause-initial position in verb-second declaratives, Bohnacker & Rosén (2008) showed that L1-Swedish learners of German had persistent problems integrating discourse-pragmatic constraints with syntax (cf. e.g. Belletti et al., 2007). Having compared native Swedish and German corpora, they found significant quantitative differences concerning the frequencies with which constituents occurred in the prefield. For instance, clause-initial subjects, particularly expletives, were more frequent in Swedish than in German, while fronted objects and certain adjuncts were much rarer. There were also qualitative differences concerning the mapping of information structure and linear word order: Swedish exhibited a stronger tendency than German to place new information, the so-called rheme, later in the clause. The L2 learners transferred these patterns from their L1 to German. Their sentences were syntactically well-formed but had Swedish-style prefield frequencies and a strong pattern of rheme later which native Germans perceive as unidiomatic, as an acceptability judgment and a rewrite-L2 texts task showed. The present study extends Bohnacker & Rosén's (2007, 2008) work in three ways. Learners of the reverse language combination (L1 German, L2 Swedish) are investigated to see whether similar phenomena also manifest themselves there. Secondly, written and oral data from highly advanced learners are included to see whether the learners' persistent problems can be overcome by long immersion in the L2. Thirdly, besides investigating theme-rheme (old vs. new information), some comments are made concerning another information-structural level, background vs. focus.

Keywords: Swedish; German; L2; information structure; expletive

1. Introduction

Interfaces, a term adopted from chemistry and electronics, are in vogue in theoretical linguistics and acquisition research today. They are areas where different components (or modules) of the linguistic computational

system interact with each other or with language-external cognitive systems, such as the conceptual-intentional and articulatory-perceptual systems. Well-known diagrams such as in (1) depict these outward-looking, grammar-external interfaces, Logical form (LF) and Phonetic form (PF) (e.g. Chomsky, 1993; Jackendoff, 2002), but not the inward-looking, grammar-internal interfaces.

(1)

Conceptual- -intentional sys- -tem	LF	Grammar (computa- -tional system & -lexicon)	PF	Articulatory- -perceptual -system
--	----	--	----	---

For second language (L2) acquisition, it has been argued that learners suffer particular problems at the interface, i.e. that they have problems integrating different kinds of grammatical knowledge, such as phonology with morphology, morphology with syntax, syntax with the lexicon, or syntax with discourse. However, there is no reason to assume that all interfaces are equally problematic – some interfaces may continue to cause problems for L2ers at advanced proficiency levels, others may not. Grammar-external interfaces have been said to be more problematic for learners than grammar-internal ones, though this is a matter of dispute (cf. White, 2009). Some researchers suggest that the interface between syntax and other cognitive systems, in particular discourse pragmatics or information structure, may be the hardest to master and that it is here where interlanguages exhibit optionality, instability and residual effects of L1 discourse strategies (e.g. Tsimpli & Sorace, 2006; Belletti *et al.*, 2007). This is reminiscent of earlier work by Carroll, von Stutterheim and collaborators on narrative production and description tasks, where advanced L2ers subtly diverged from native speakers in the ways they employed word order to structure oral and written texts, retaining core conceptual patterns and principles of their L1 (Carroll & Lambert, 2003; Carroll & von Stutterheim, 2003).

2. The interface of syntax and discourse-pragmatics

The syntax of a language is commonly described as a set of rules, parameters or constraints on which orderings are possible irrespective of context. In a particular context, certain constituent orderings (e.g. preverbal and postverbal subjects) may be more likely or more felicitous than others but this variation is typically not ascribed to pure syntax, but to semantic and discourse-pragmatic factors. Discourse pragmatics covers many phenomena including politeness marking and language choice in multilingual contexts, but for present purposes, another area of discourse pragmatics is more relevant, namely the way how speakers/writers of a particular language organise and present information. Such information management can be studied at a

global or text level (e.g. Halliday & Hasan, 1976; von Steutterheim, 1997) or at a local (i.e. utterance or clause) level. It is this latter, local level that interests us here – language-specific information-structural influence on constituent ordering.

Information structure concerns the division of information into more or less salient or relevant and its packaging and presentation with the help of linguistic structure (e.g. Chafe, 1976; Prince, 1981). The information structure of an utterance depends on the context the utterance occurs in. Information structure is sometimes reduced to two notions, topic and focus, but the central European tradition regards these two as a conflation of different levels of information structure and therefore employs notions beyond topic and focus, at several distinct levels (e.g. Jacobs, 1984; Krifka, 2007). Here it is worth reminding ourselves that the way we view information structure depends on our model of communication. A widely known model is Bühler's (1934) communicative triangle (2), where content (*Gegenstände und Sachverhalte*) is communicated by the sender (*Sender*) to the receiver (*Empfänger*).

(2) Gegenstände und Sachverhalte



- (3) a. topic vs. comment information structure
 b. theme vs. rheme information structure
 c. background vs. focus information structure

Bühler's triangle yields three separate information-structural levels, (3a) topic vs. comment, (3b) theme vs. rheme, and (3c) background vs. focus. At the level of content (*Gegenstände und Sachverhalte*), information is typically divided into topic and comment. The topic of a sentence is the point of departure and it provides the referential frame with respect to which the predication is evaluated. Information *about* the topic is encoded in the comment (cf. Reinhart, 1982 and her file card image; Lambrecht, 1994; Krifka, 2007). At another level, that of the receiver (*Empfänger*), information is structured into theme and rheme, in the terminology of Ammann (1928) and the Prague School (e.g. Daneš, 1970; Beneš, 1971; cf. also Prince, 1981). *Theme* stands for what the speaker/writer assumes the listener/reader to know; it is old, maintained or given information in the sense that it has previously been explicitly mentioned or is inferable with recourse to the linguistic discourse or the discourse situation. *Rheme* stands for what the speaker assumes to be new information for the hearer¹. Finally, at the level of

¹ Dividing the clause into theme and rheme is not always straightforward, as clauses may contain several thematic elements, and some contain none but are informationally all-new. Thematicity/givenness may also be viewed as a graded property, where recency of mention and other factors influence how accessible a thematic/given element is.

Sender, information can also be divided into background and focus, depending on the speaker's choice of demoting (background) and highlighting information (focus) (cf. Jacobs, 1984; Krifka, 2007). A tripartite pragmatic approach to information structure such as the one in (3) allows the different information-structural levels to correlate, and they often do. For instance, focused information is frequently encoded by the same expression as new information and tends to occur in the part of the utterance that contains the comment. Also, topic (point of departure) and theme (old information) frequently coincide in an utterance. However, not all topics are themes, and not all themes are topics. Likewise, what is focused in an utterance need not be new information, since focus and rheme relate to different information-structural levels. Importantly then, a correlation between the levels is not at all mandatory. In section 3, these notions are related to the German and Swedish prefield, but before doing so, I will make some general comments on the relation (and the interface) between syntax and discourse pragmatics.

How to view and formalise the relation between syntactic form and discourse function is much debated, and my understanding of the literature leads me to distinguish three major lines of approach. One is the in essence functionalist view that the grammatical form directly follows from the communicative function of a sentence (cf. Kuno, 1987). Another line of approach is the 'traditional' generative view that syntax is autonomous and discourse function is external to syntax (e.g. Chomsky, 1965; Prince, 1981, 1998; Fanselow, 2007; Féry, 2007). Prince (1998:281) puts it as follows: "the relation between syntactic form and discourse function is no less arbitrary than, say, the relation between phonological form and lexical meaning". Structural possibilities are provided by the grammar *independently* of discourse pragmatics, and discourse-pragmatic notions do not play a role in the identification of syntactic slots or categories, nor in the triggering of syntactic operations. A multitude of grammatical devices (phonological, morphological and lexical markers, syntactic structures and surface positions) may be employed to support different discourse functions, but there is great crosslinguistic variation and particular discourse functions do not invariably correlate with any grammatical reflex, according to the autonomous-syntax view (Prince, 1998; Féry, 2007). Any mapping between language-specific form and pragmatic function can thus only be indirect and takes place not in syntax but in separate cognitive components (Lambrecht, 1994; Costa, 2004; Neeleman & van de Koot, 2008)².

This view contrasts with a third approach, where information-structural notions such as topic, focus, givenness are incorporated into formal theories of syntax via an articulated hierarchy of functional projections and corre-

² Due to effects of pressure on planning and language production in real time, there may also be processing explanations for correlating linguistic-form/discourse-function tendencies (e.g. Arnold *et al.*, 2000).

sponding ‘syntactic’ features. Topic and focus are regarded as morphological notions in some abstract sense, attracting relevant syntactic constituents to a specifier in order to check a morphological feature (e.g. Rizzi, 1997; Belletti, 2004). Particular discourse-related interpretations are licensed in the specifier of a designated functional projection – in other words, “syntactic positions—ultimately word order—directly affect aspects of the interpretation, which can thus be read off the syntactic configuration” (Belletti *et al.*, 2007:659). This ‘cartographic’ approach may be attractive for languages where there are discrete morphological markers for particular discourse functions but becomes rather abstract for languages such as German or Swedish which neither have such overt morphological markers nor fixed information-structural slots. Whilst cartographic approaches are gaining in popularity, proponents do not agree on the details of this proliferated phrase structure. As Benincà & Poletto (2004:52) put it, “there is no limit, in our view, as to how many of these projections there will ultimately be”. This may raise questions about learnability and concerns that formal theories of syntax try to account for phenomena that would better be handled in semantic, pragmatic or processing terms.

There is thus little agreement on whether and how much discourse pragmatics should be represented in the syntax, and I won’t be taking a stand on this matter here. However, I would like to point out that depending on which line of approach is chosen, the locus or type of the interface in one’s model of language knowledge may change. In the generative tradition, linguistic competence is mentally represented by means of an abstract linguistic system, the grammar. In this grammar, different components or modules interface with each other grammar-internally, and they also interface with other, grammar-external domains such as the conceptual-intentional system. A ‘discourse-free’ syntax approach thus necessarily involves an external interface with an interpretive module. A ‘discourse-laden’ syntax à la Rizzi or Belletti strives to treat discourse-pragmatic notions essentially as syntactic and as part of the computational system of the grammar. But if they are part of the computational system this suggests a grammar-*internal* interface (notwithstanding the existence of a grammar-external interface with an interpretive module). Recent L2 research points fairly consistently to learner problems associated with phenomena that involve a relationship between syntax and discourse pragmatics. In the past, generative linguists have often relegated these to domains outside the grammar proper, e.g. to pragmatics, stylistics or rhetoric (Liceras, 1988). Nowadays, they are described as interface problems. Alternatively, they are argued to arise within the computational system itself – as representational deficits in functional categories or features associated with these categories. But it seems to me that there is no theory-neutral answer to the question whether L2ers have greater problems at grammar-external interfaces than at grammar-internal interfaces, and to whether the problems are pragmatic or grammatical in nature, since the answer very much depends on the formal theory of syntax chosen.

3. The prefield at the interface of syntax and information structure

Both Swedish and German adhere to the verb-second (V2) constraint that requires the finite verb in declaratives to be the second constituent. In non-subject-initial main clauses, inversion of the subject and the verb is required, and V3 is generally ungrammatical. The position to the left of the finite verb is called the ‘prefield’ (e.g. Reis, 1980). In principle, this position can contain virtually any constituent in German and Swedish irrespective of syntactic category, complexity and semantic function, some modal particles excluded (e.g. Zifonun *et al.*, 1997; Teleman *et al.*, 1999).

(4)	Prefield	Finite verb	Remainder of declarative clause
a. Swe.	Nu	har	väl alla fått ett ex.
	now	have	well all got a copy
b. Ger.	Jetzt	haben	ja wohl alle eins.
	now	have	yes well all one
			‘I suppose everyone’s got a copy now.’

On the view that a description of the syntax of a language should aim for specifying the possible strings, regardless of the likelihood that such strings will sound felicitous in a particular discourse context (e.g. Prince, 1998), we can say that the syntactic constraints on the prefield and on how to start a V2 declarative clause in the two languages are the same. Not surprisingly, Swedish and German as well as other Germanic V2 languages are assumed to behave alike concerning the function and frequency of prefield constituents, with a distribution of 70% or 60% subject-initial vs. 30% or 40% non-subject-initial, though such figures are usually not based on corpora counts. However, when Christina Rosén and I surveyed existing text corpora we found that V2 languages may differ substantially in the way they make use of the prefield, both quantitatively and qualitatively (Bohnacker & Rosén, 2007). Subject-initial clauses are consistently more frequent in Swedish than in German. The corpus data we collected ourselves confirmed this; as shown in Fig. 1, Swedish has a stronger subject-initial preference (73%) than German (50%); objects are fronted more often in German (7%) than in Swedish (3%), and adverbials other than temporal and locational are fronted more frequently in German (18%) than in Swedish (6%)³.

³ These differences are statistically significant for subjects and expletives ($\chi^2=75.797$, $p<0.001$), objects ($\chi^2=15.216$, $p<0.001$) and other adverbials ($\chi^2=58.951$, $p<0.001$), but not for temporal and locational adverbials. Figure 1 only compares informal letters in order to avoid any potential confounding effects that different text types might cause.

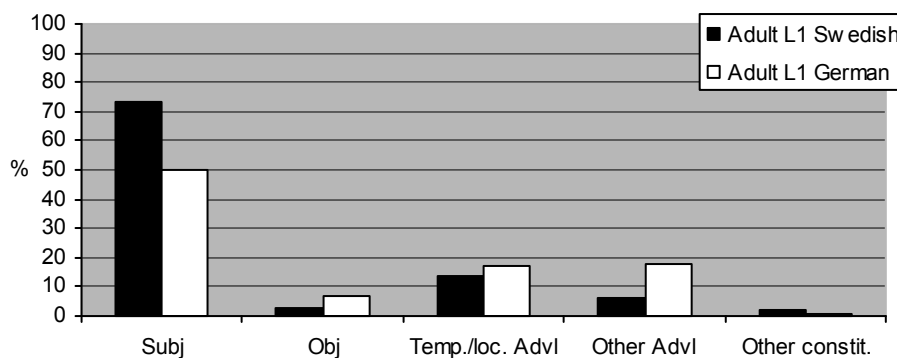


Figure 1: Overt constituents in the prefield, written L1 data (Bohnacker & Rosén, 2008)

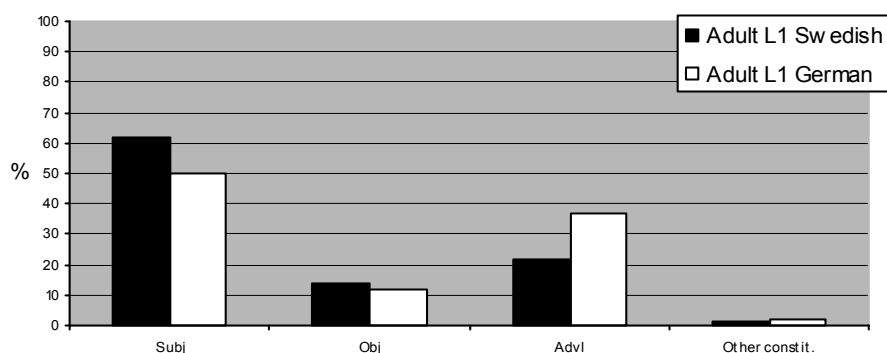


Figure 2: Overt constituents in the prefield, oral L1 data (Jørgensen, 1976; Bohnacker, unpublished)

A similar asymmetry can be found for informal speech (Fig. 2). I show this for a new corpus of colloquial spoken German that I collected and transcribed myself (Bohnacker, unpublished, see Section 4) and spoken Swedish corpus data from the Talsyntax project at Lund University (Jørgensen, 1976). Swedish again has a stronger subject-initial preference (62%) than German (50%) and fronts adverbials less often (22%) than German (37%), though there appears to be no pronounced difference for fronted objects (14%, 12%) in the spoken data.

These differences in frequency between Swedish and German led Rosén and myself to investigate the prefield in the two languages more closely. We were struck by the fact that Swedish speakers mostly placed phonologically light elements in clause-initial position, elements that had low or no informational value. Concerning subjects, it was particularly interesting to see that expletive *det* 'it' in the prefield was much more frequent in Swedish than expletive *es* 'it' in German. In Rosén's (2006) corpora of informal letters,

22% of all subject-initial sentences beginning with an expletive in Swedish, but only 11% do so in German, as shown in Table 1. This difference is significant ($\chi^2 = 48.00$, $p < 0.001$).

	Expletives out of all subjects	Expletives out of all overt prefield constituents
Adult L1 Swedish	22% (85/388)	16% (85/535)
Adult L1 German	11% (66/587)	6% (66/1173)

Table 1: Expletive subjects in the prefield, written data, informal letters (Bohnacker & Rosén, 2008:520)

Interestingly, the same asymmetry concerning expletive subjects can be found for informal speech (Table 2). I compare two corpora of spoken Swedish (Jørgensen, 1976) with new corpus data from spoken colloquial German. 16% and 19% of all subject-initial declaratives in the oral Swedish data start with an expletive, but in the oral German data only 3% do so. The difference between the Swedish and German corpora is again significant ($\chi^2 = 221.08$, $p < 0.001$).

	Expletives out of all subjects	Expletives out of all overt prefield constituents
Adult L1 Swedish, Jørgensen A	16% (99/632)	10% (99/979)
Adult L1 Swedish, Jørgensen B	19% (578/3068)	13% (578/4610)
Adult L1 German, Bohnacker corpus	3% (39/1190)	2% (39/2294)

Table 2: Expletive subjects in the prefield, informal oral data (Jørgensen, 1976; Bohnacker, unpublished)

Constituents other than subjects in the prefield also deserve to be properly investigated, but due to space constraints I cannot do so here but only briefly comment on objects. The definite inanimate pronoun *det* 'it/that' is by far the most common fronted object in Swedish and more frequent in the prefield than its German equivalent *das* 'it/that'. German speakers front a wider range of objects, both lexical and pronominal. For instance, in the informal written L1 texts collected by Rosén, *det* makes up 82% of all fronted object pronouns, but *das* only 24% (Rosén, 2006:99-102). A similar asymmetry seems to hold for the oral data (Bohnacker, in progress).

The precise percentages of elements in the prefield (cf. Figs. 1-2, Tab. 1-2) may be different for corpora of other text types, but importantly, there is a clear asymmetry between German and Swedish when keeping genre constant. I will argue that this is likely to be due to different tendencies concerning the mapping of syntax and information structure.

The prefield is particularly important for communication as it anchors the clause in discourse. At the inter-sentential level, the prefield contributes to textual coherence by linking up with preceding discourse; at the intra-sentential level, it often establishes the topic identified by the speaker, about which s/he then provides information (cf. Section 2). Yet this does not mean that the prefield is a reserved topic slot; topics can also occur elsewhere and non-topics can occur in the prefield as well. As regards theme and rheme, Swedish and German have a tendency towards given before new, a tendency attested for many languages which may have to do with ease of online processing. This tendency, coupled with the V2 constraint of the two languages, gives rise to clauses where the prefield contains an element of low informational value. New information, the rheme, is usually provided later, after the finite verb (cf. Daneš, 1970; Beneš, 1971; Teleman *et al.*, 1999:53-64). Alternatively, the prefield can host a focused element, coded prosodically via stress and pitch contours, to be contrasted with other members of some evoked set of alternatives (cf. e.g. Zifonun *et al.*, 1997; Prince, 1998). These observations are not new, and so far they suggest that the two languages are information-structurally similar. There are certain word order tendencies, but little evidence for a *direct* impact of information structure on Swedish and German syntax. Neither language appears to have a fixed slot for elements with a particular information-structural function, in contrast to what has sometimes been argued for other languages, such as a preverbal focus position for Korean or Hungarian (but see Féry, 2007).

Nevertheless, Bohnacker & Rosén's (2007, 2008) comparative corpus data (as well as an acceptability judgment task not reported on here) suggest that V2 languages may differ in the way they make use of the prefield, both quantitatively and qualitatively. We found that Swedish has a stronger tendency than German to keep informationally new (i.e. rhematic) material out of the clause-initial position and instead place it further to the right, i.e. post-verbally. This can be achieved by filling the prefield with given (i.e. thematic) information, or with an element of no informational value, such as an expletive subject, or by leaving the prefield empty, as in V1 declaratives (not discussed here). We might thus say that Swedish linear syntax more faithfully follows the information-structural principle of *Rheme later*, schematised in (5). The examples in (6) illustrate this: Swedish disprefers clause-initial rhematic subjects; rhematic subjects (e.g. *en massa folk* 'lots of people') are nearly always postverbal and the prefield is filled by an expletive subject. No such tendency can be discerned for German – it is perfectly acceptable to start off with a rhematic subject (e.g. *ne Menge Leute* 'lots of people' in (6'a)) and more common in our corpus data than filling the prefield with an expletive (6'b).

- (5) Rheme later

<i>Prefield</i>	<i>Finite verb</i>	<i>Middle field etc.</i>
Expletive or given information		New information

- (6) A: Anything happened this morning?
 B: a. **Det** har ringt en massa folk till dig. (Swedish preferred)
 EXPL has called a lot people to you
 b. **En massa folk** har ringt till dig. (Swedish dispreferred)
 ‘Lots of people have been calling you.’
- (6’) a. **Ne Menge Leute** haben dich angerufen. (German preferred)
 a lot people have you called
 b. **Es** haben dich ne Menge Leute angerufen. (German dispreferred)

Likewise, clause-initial rhematic objects are rare in Swedish – if fronted, objects are nearly always themes (old information) and simply consist of an anaphoric definite pronoun, especially *det* ‘it/that’, as in (7). Its German equivalent *das* appears in the prefield too, but German speakers regularly front a wider range of pronominal and lexical objects, such as *Kissen und einen blauen Flickenteppich* ‘cushions and a blue rug’ (8). Swedish speakers would instead start with a thematic subject (*jag* ‘I’) and postpone the object *kuddar och en blå trasmatta* (8’).

- (7) A: We’ve got a special offer today – vegetable lasagne for 3.99.
 B: **Det** tar vi.
 that take we
 ‘We’ll have that.’
- (8) Gestern war ich bei Ikea und hab zwei Regale besorgt. (German)
 yesterday was I at Ikea and have two shelves got
Kissen und einen blauen Flickenteppich hab ich auch gekauft.
 cushions and a blue rug have I also bought
 ‘Yesterday, I went to Ikea and got two shelves. I also bought cushions and a blue rug.’
- (8’) a. [?]**Kuddar och en blå trasmatta** köpte jag också.
 (Swedish dispreferred)
 Cushions and a blue rug bought I also
 b. Jag köpte också kuddar och en blå trasmatta. (Swedish preferred)

In addition to these divergent tendencies concerning given vs. new (theme-rheme), there may also be different word order tendencies at another information-structural level (background-focus), when fronted objects are considered. As mentioned above, Rosén (2006) found that inanimate *det* ‘it/that’ made up the bulk of fronted pronominal objects. Yet why would Swedish front pronominal objects other than *det* less frequently than German? Such a difference cannot straightforwardly be accounted for by saying

that Rheme later is stronger in Swedish. Pronouns in both languages typically encode old information (the theme) as they refer back to a textually accessible antecedent. For instance, both personal object pronouns *die* ‘her’ and *henne* ‘her’ in (9)-(9’) function as themes, as they refer back to the previously mentioned Louisa. Why is it then that a fronted animate object is dispreferred in Swedish (9’a)?

- (9) A: And when’re you gonna ask Louisa?
 B: a. **Die** hab ich schon gefragt. (German preferred)
 her have I already asked
 b. Ich hab die schon gefragt. (German dispreferred)
 ‘I’ve already asked her.’
- (9’) a. ?**Henne** har jag redan frågat. (Swedish dispreferred)
 her have I already asked
 b. Jag har redan frågat henne. (Swedish preferred)

The answer I believe lies at the level of background-focus. Animate personal object pronouns in the Swedish prefield are not impossible per se, but they often carry stress, thus inducing an interpretation of narrow/minimal focus. This focus introduces a set of alternatives that contrast with the meaning of the utterance. For a native speaker of Swedish, fronted *henne* in (9’) would evoke a situation where the speaker contrasts having asked Louisa with not having asked one or several other persons. As no such narrow/minimal focus reading is intended, the object pronoun remains both unstressed and unfronted (9’b)⁴. Hazarding a guess, I would expect that narrow/minimal focus on pronominal objects is not very common in text corpora, and if this turns out to be true, it could explain why fronted personal object pronouns are infrequent in the Swedish data. In German, there is nothing wrong with unstressed personal object pronouns in the prefield, so the numbers of fronted personal object pronouns are higher. Obviously, these speculations will need to be investigated more thoroughly in future work.

⁴ Aafke Hulk (21 June 2010) informs me that fronted Dutch animate personal pronouns receive a focus reading too.

4. Informants, data collection and method

The L2 learners were eight adult native speakers of German. They had all had a monolingual childhood in Germany or Austria. At school they had had 7-9 years of English as a foreign language (from age 10/11), and some years of Latin or French. As regards Swedish, they were adult learners; none had been exposed to Swedish before the age of 20. At the time of the study, the informants were long-term residents of Sweden and they used Swedish every day, in the workplace, with friends, and/or at home. They were university graduates in their early twenties to late thirties, employed at schools, universities and with the local council, as teachers, researchers, cleaners, and therapists. Whilst all had been exposed to classroom Swedish, most of their acquisition was naturalistic⁵. The learners were advanced in the sense that they were communicating fluently and had passed the respective Swedish university-entry language proficiency exams (Rikstestet/TISUS) before data collection started⁶. The learners stated that they felt at ease when speaking Swedish but less confident when writing the language.

The data from five of these informants are longitudinal and were collected at three-year intervals; the remaining three were studied once⁷. All the data are naturalistic production, spoken and written. The oral data consist of a 45-min recording of the informant narrating events of their life in conversation with an experimenter, and a 45-min recording of the informant giving a class/seminar at the workplace in the absence of experimenters. Each recording consists of 5000-7000 words and contains both dialogue and monological passages. Additionally, the informants each supplied 5000 words of unedited written text (informal emails/letters). Word order and constituent type were coded and classified by hand (Bohnacker, 2007). Oral and written data are investigated separately, so as not to mask potential information-structural differences between the two modes. However, I have collapsed oral narrative and oral teaching into one informal oral category, as there were no substantial differences between them. The L2 data were originally collected for a study on verb placement and verb particle constructions

⁵ Three had attended classes in Swedish as a foreign language in Germany (2hrs/week for one year) and began to work immediately upon arrival in Sweden, without taking further classes. The other learners had no previous knowledge of Swedish before arriving in Sweden. They attended Swedish classes for immigrants (4-10hrs/week) for one year, after which acquisition continued untutored.

⁶ TISUS (Test In Swedish for University Studies) is a two-day examination, testing reading comprehension, written composition, and oral communication in an interview with two examiners.

⁷ From Stefanie, data were collected 3 years after arrival, from Stella, Nicole and Ellen after 3 and 6 years, from Ulrike and Steffen after 3, 6 and 9 years, and from Emma and Dirk after 15 years in Sweden.

(Bohnacker, 2007), but can also be used to study the prefield of declaratives. There are 11,555 declaratives or instances of a filled or potentially filled prefield, 4044 for the written data and 7511 for the oral data. The learners placed the finite verb in a targetlike manner, with only 0.026% violations of the V2 constraint (Bohnacker, 2007: 24-26).

The native controls are adults and roughly of the same age group as the L2 learners. L1 text types have been matched as closely as possible with those of the L2ers. The written L1 corpora comprise compositions (informal letters, summaries) by 70 native speakers of German (28,5000 words) and 80 native Swedish speakers (17,500 words), from 1999-2005 (see Rosén, 2006.)

The oral L1 data consist of conversations between native speakers that include both dialogue and more monological narrative passages. Here I am using a new, previously unpublished, corpus of oral native German, which contains 25,300 words of colloquial South German dialogue (Bohnacker). Three female informants aged 25-35 and one aged 60, all from the greater Ulm area, were recorded for 3 hours in one-on-one conversations with a local experimenter. The recordings were made between 1994 and 2000 and transcribed by myself.

For oral native Swedish, I did not have access to an original corpus, but refer to Jörgensen's (1976) corpus study of recordings made in 1968. This includes (i) 32 informal interviews of 30-to-45-year-old employees on the topic of immigrants and immigration (8-9 hours of recording), and (ii) conversations and discussions between native Swedish academics (8 informants, 3 hours of recording). The interviews comprise 45,000 words, the conversations and discussions 11,200 words.

5. Results: How German-speaking learners of Swedish make use of the prefield

5.1. Subjects and expletives

The overall frequencies of subject-initial clauses out of all V2 declaratives in the L2 Swedish productions are closer to native German (50% written, 50% oral) than to native Swedish (73% written, 62% oral). As shown by the black bars in Figs. 3 and 4, in the L2 writing, the percentages of subject-initial declaratives hover around 37%-50%, and around 49%-60% in the oral L2 data⁸. The white bars represent expletive-initial clauses and will be discussed shortly. For ease of exposition, the data have been aggregated for the learners at 3 vs. 6, 9, and 15 years.

⁸ There are more subject-initial clauses (60%) in the learners' oral data at 3 years than at any other data point. This is due to a preponderance of first person singular subjects (*jag* 'I') in some of the data files because of the way the interviewer interacted with the informant.

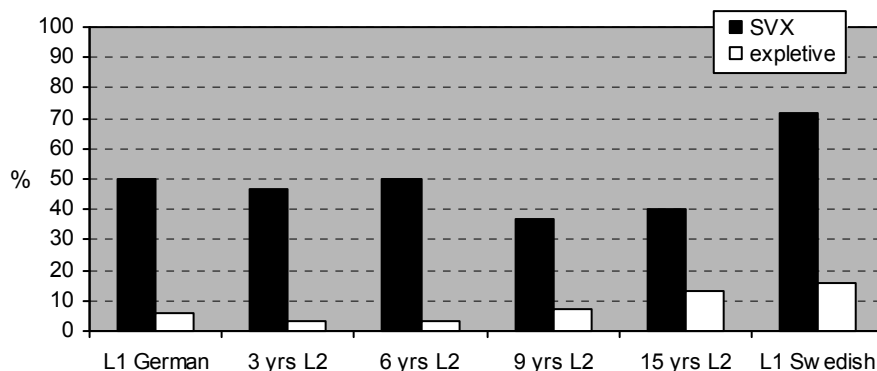


Figure 3: Prefield subjects and expletives in L1 German, L2 Swedish and L1 Swedish, informal written data (letters)

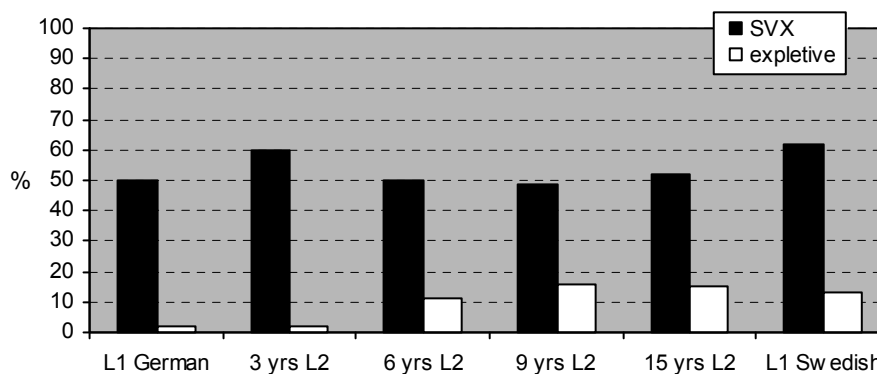


Figure 4: Prefield subjects and expletives in L1 German, L2 Swedish and L1 Swedish, informal speech

When we investigate the informants' subject-initial declaratives more closely, developmental tendencies emerge. At 3 years, informationally new and phonologically heavy subjects regularly occur in the prefield, and in some speakers also at 6 years and 9 years (e.g. (10)-(11)). Whilst not ungrammatical, these heavy clause-initial subjects are unusual in L1 Swedish, where one would preferably start off with a light expletive subject and place the rhematic subject postverbally, as in (10') and (11')⁹.

⁹ Moreover, the learners' choice of lexical items, grammatical gender and inflectional morphology sometimes differs from native Swedish. Such nontarget features will not be commented on here.

- (10) Tack för korten. **En riktigt fin sjungvecka** var det med er!
 Thanks for pictures-the a really nice sing-week was it with you
 ‘Thanks for the pictures. It’s been a really nice singing week with you.’
 (L2er Ellen, 3 years, written)
- (10’) Native preferred: **Det** har varit en riktigt fin körvecka med er.
 EXPL has been a really nice choir-week with you
- (11) ‘Here you can see at it increases with age, but you can’t see it so well here.’
 en **en bättre diagram** är på sida 67.
 but a better chart is on page 67
 ‘But there’s a better chart on page 67.’
 (L2er Ulrike, 3 years, oral, teaching)
- (11’) Native preferred: Men **det** finns ett bättre diagram på sidan 67.
 but EXPL is a better chart on page-the 67

There is a clear trend in the learner data concerning the percentage of clause-initial expletive *det*, plotted as white bars in Figs. 3 and 4. At 3 years, the learners rarely produce clause-initial expletive *det* (2% written, 3% oral), which is significantly lower than the native Swedish speakers (16% written, 13% oral)¹⁰, but similar to native German (6% written, 2% oral). (Exact figures are provided in Tab. 3-4.) As the learners produce a large number of declaratives (3904) at this point, this is unlikely to be a sampling artefact but can be interpreted as transfer of the L1 German pattern to Swedish. From 6 years onwards, the proportion of clause-initial *det* goes up, though this is evident at first in the oral data only. Over the years, clause-initial expletive *det* increases six to ten times, from 3% at 3 years to 7% and 13% (written) and 16% and 15% (oral) at 9 and 15 years, respectively. I suggest that this change is indicative of the learners’ growing awareness of the Swedish frequency distributions and information-structural patterns, with a strong preference for rhematic information being placed later in the clause.

¹⁰ The difference between the L2ers at 3 years and the native Swedish speakers is significant both for the written condition ($\chi^2 = 117.01$, $p < 0.001$) and the oral condition ($\chi^2 = 194.71$, $p < 0.001$).

In Sweden	Expletives out of all subjects	Expletives out of overt prefield constituents
3 yrs, 6 learners	5% (35/648)	3% (35/1380)
6 yrs, 5 learners	5% (29/564)	3% (29/1121)
9 yrs, 2 learners	19% (46/240)	7% (46/654)
15 yrs, 2 learners	32% (71/224)	13% (71/558)

Table 3: L2 Expletive subjects in the prefield, informal written L2 data

In Sweden	Expletives out of all subjects	Expletives out of overt prefield constituents
3 yrs, 6 learners	3% (45/1306)	2% (45/2179)
6 yrs, 5 learners	22% (210/962)	11% (210/1917)
9 yrs, 2 learners	33% (185/561)	16% (185/1158)
15 yrs, 2 learners	30% (178/593)	15% (178/1146)

Table 4: L2 Expletive subjects in the prefield, informal oral L2 data

5.2. Objects

In general, the learners front more objects than native Swedish speakers do, as shown by the black bars in Figs. 5-6. Due to space constraints, I have to be very brief here. Whilst there are few fronted *det* 'it/that' at 3 years, the proportion of *det* (white bars) out of all fronted objects (black bars) increases over time, more so in speech than in writing.

Besides *det*, the learners front a variety of objects, both pronominal and lexical, including ones that are informationally new, as in (12)-(13). Common at first in both speech and writing at 3 years, in later years, these are mainly found in the L2 writing. Native Swedish speakers perceive such sentence openings as unidiomatic, heavy, stilted, old-fashioned and un-Swedish and would instead start off with a light subject pronoun (*jag* 'I', *du* 'you') and place the rhematic object postverbally, cf. (12'), (13').

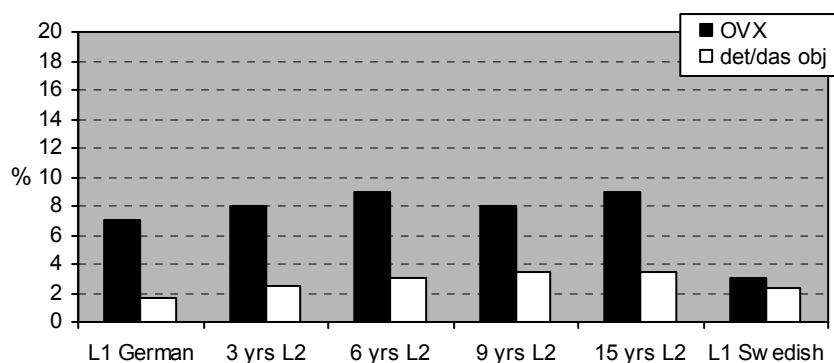


Figure 5: Prefield objects and object *det/das* in L1 German, L2 Swedish, L1 Swedish, informal writing

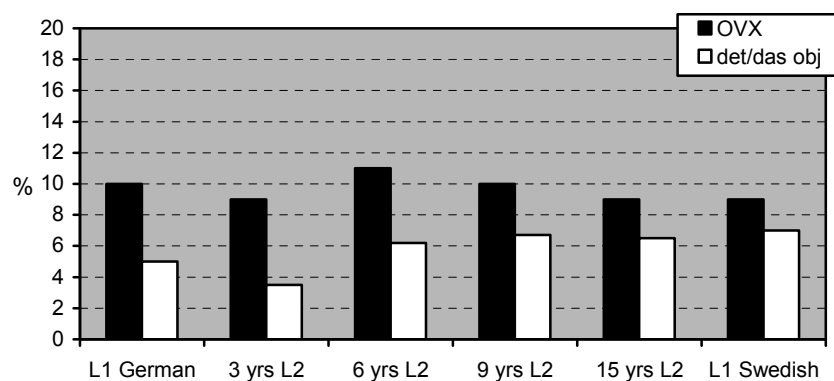


Figure 6: Prefield objects and object *det/das* in L1 German, L2 Swedish, L1 Swedish, informal speech

- (12) I lördags har jag varit på IKEA och köpt två bokhyllor.
 on Saturday have I been at IKEA and bought two bookcases
 ‘On Saturday, I went to IKEA and bought two bookcases.’
En blå trasmatta och en kudde har jag också köpt.
 a blue rug and a cushion have I also bought
 ‘I also bought a blue rug and a cushion.’
 (L2 Ulrike, 3 years, written)

- (12’) Native: Jag köpte även en blå trasmatta och en kudde.
 I bought also a blue rug and a cushion

- (13) Hoppas att du har haft roligt igår. (L2 Nicole, 3 years, written)
 hope that you have had fun yesterday
 ‘I hope you had fun yesterday.’
En cykeltur eller en längre tur till havet har du gjort, kanske?
 a bike-trip or a longer trip to sea-the haveyou made perhaps
 ‘You went for a bike ride or a longer trip to the seaside perhaps.’

- (13’) Native: Du (kanske) gjorde en cykeltur eller en längre tur till
 havet (kanske)?
 you (maybe)made a bike-trip or a longer trip to
 sea-the (maybe)

Also of interest is the occurrence of fronted pronominal objects other than *det*, in particular animate personal pronouns. As noted in Section 3, fronted objects carry emphatic/contrastive stress and minimal/narrow focus, inducing an interpretation of contrast in native Swedish. Consider now (14), where Nicole fronts the (unstressed) prepositional object *till honom* ‘to him’, referring back to *din chef* ‘your boss’ mentioned earlier. For a native speaker of Swedish, fronted *till honom* evokes a situation where having spoken to the boss is contrasted with not having spoken to one or several other persons. No such interpretation appears to be intended by Nicole, and there is no stress on *till honom*, which makes the fronted object inappropriate and the sentence “un-Swedish”. The same goes for the objects in (15)-(16). I suggest that they are due to L1 transfer, since in the learners’ native German, such fronted pronominal objects would be acceptable and simply interpreted as neutral, given information (cf. (15’), (16’)).

After 3 years of immersion in a Swedish-dominant environment, the learners still seem unaware of the subtle interpretive effects of placing objects in the left periphery of the clause. This is reminiscent of Camacho’s (1999) findings on L1 Peruvian Quechua speakers acquiring L2 Spanish. In Quechua, left-peripheral objects without clitics are information-structurally “neutral” (i.e. they do not evoke narrow/contrastive focus). Transferring this neutral L1 pattern to the L2, the learners produced objects without clitics in the left periphery in their Spanish, unaware that native Spanish listeners assign contrastive focus to them.

- (14) Context: About doing a training course abroad and how to go about telling one’s boss.
 I: Å när ska du berätta de(t) för din chef?
 and when shall you tell it for your boss
 ‘And when’re you gonna tell your boss?’
 N: **Till honom** har ja(g) redan sagt det, men det blir kanske inte av i
 to him haveI already said it but it become maybe not off in
 alla fall.
 any case
 ‘I’ve already told him, but it might not happen anyway.’
 (L2 Nicole, 3 yrs, oral)

- (14') a. German: {**Dem/ihm**} hab ich das schon gesagt.
 him-DAT have I it already said
 b. Swedish: Jag har redan {sagt/berättat} det {till/för} honom.
 I have already said/told it to/for him
- (15) Det gör inget att du inte kan nu. **Oss** ska sen pappa hjälpa.
 it does not that you not can now us will then dad help
 'It doesn't matter if you can't come (and help) now. Dad will help us later.'
 (L2 Stella, 3 years, written)
- (15') a. German: Uns wird dann der Vater helfen.
 us will then the father help
 b. Swedish: Pappa kommer-att hjälpa oss sen.
 dad will help us later
- (16) 'I haven't heard anything from XY anymore.'
Henne har jag också skickat artikeln.
 her have I also sent article-the
 'I sent the article to her as well.' (L2 Ulrike, 6 years, written)
- (16') a. German: {**Der/ihr**} hab ich den Artikel auch geschickt.
 her-DAT have I the article also sent
 b. Swedish: Jag har också skickat artikeln till henne.
 I have also sent article-the to her

6. Conclusion

In this paper, I have suggested that German-L1 advanced learners of L2 Swedish exhibit German-style frequency patterns in the prefield and German-style information-structural patterns which differ from native Swedish. This is particularly striking for the earliest data presented here, i.e. naturalistic production after 3 years. Even though the data are only longitudinal for some individuals and otherwise cross-sectional, they are suggestive of development towards the target: by 6 and 9 years, the rates of clause-initial expletive subjects and lightweight given elements (pronominal object *det*) go up, unidiomatic clause-initial rhematic elements decrease, and clause-initial animate personal object pronouns with an unwarranted narrow focus interpretation are less common.

The development towards the target sets in earlier and more forcefully in L2 speech than in writing. This may be a surprising result since in other studies (of other phenomena) L2ers typically do better in unsped writing tasks because they have time to monitor. Our learners may also have monitored their writing for a number of things, yet it is unlikely that they monitored for differences concerning the interaction of information-structure and

word order, as they were probably not even aware that Swedish differs from German in this regard. The word order preferences under discussion are not either/or, grammatical/ungrammatical, they are not readily accessible to introspection and largely ignored in descriptive grammars, teacher training and language teaching materials. Residual L1 patterns might however show more strongly in the L2 writing because writing is what the learners of Swedish do the least – their acquisition and immersion has been mainly oral and aural (cf. Section 4). The learners do better than Bohnacker & Rosén's (2008) learners of the reverse language combination (L1 Swedish, L2 German). This is likely to be an effect of increased exposure (input): 3 or more years in a dominant L2 environment vis-à-vis foreign language classroom learning in Bohnacker & Rosén's study.

As for my initial discussion of interfaces in Sections 1 and 2, the learners in the present study, like those of other recent studies (e.g. Lozano, 2006; Tsimpli & Sorace, 2006; Belletti *et al.*, 2007), master pure syntax well before they master the appropriate discourse-pragmatic use of that syntax. One might want to capture this with a cartographic approach, where the language-specific information-structural differences and L2 problems would essentially be treated as syntactic and grammar-*internal* and located inside the computational system (e.g. Belletti *et al.*, 2007). Or one could describe the results with a discourse-free theory of syntax, where V2 is identical across languages, but where information-structural differences between Swedish and German and ensuing L2 problems are treated as pragmatic and outside of grammar and located at an *external* interface with the conceptual-intentional system. To my mind however, determining internal and external interface problems in L2 acquisition seems to have somewhat too much to do with one's predilection for a particular formal theory of syntax, and not only with the empirical learner data. So I won't take a stand on the interface issue here.

References

- Ammann, H. (1928). *Die menschliche Rede. Sprachphilosophische Untersuchungen 2. Teil: Der Satz*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Arnold, J. E., A. Losongco, T. Wasow & R. Ginstrom (2000). Heaviness vs. newness: The effects of structural complexity and discourse structure on constituent ordering. *Language* 76, pp. 28-55.
- Belletti, A., E. Bennati & A. Sorace (2007). Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language and Linguistic Theory* 25, pp. 657-689.
- Benincà, P. & C. Poletto (2004). Topic, focus, and V2: Defining the CP sublayers. In L. Rizzi (ed.) *The structure of CP and IP: The cartography of syntactic structures*. Vol. 2. Oxford: Oxford University Press, pp. 52-75.

- Beneš, E. (1971). Die Besetzung der ersten Position im deutschen Aussagesatz. *Fragen der strukturellen Syntax und der kontrastiven Grammatik*. Düsseldorf: Schwann, pp. 160-182.
- Bohnacker, U. (2007). On the “vulnerability” of syntactic domains in Swedish and German. *Language Acquisition* 14, pp. 1-43.
- Bohnacker, U. & C. Rosén (2007). How to start a declarative V2 clause: Transfer of syntax or information structure in L2 German. *Nordlyd* 34 (3), pp. 29-56. <http://www.ub.uit.no/baser/nordlyd>
- Bohnacker, U. & C. Rosén (2008). The clause-initial position in L2 German declaratives: Transfer of information structure. *Studies in Second Language Acquisition* 30, pp. 511-538.
- Bühler, K. (1934/1978). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Frankfurt: Ullstein.
- Camacho, J. (1999). From SOV to SVO. *Second Language Research* 15, pp. 115-132.
- Carroll, M. & M. Lambert (2003). Information structure in narratives and the role of grammaticised knowledge. In C. Dimroth & M. Starren (eds.) *Information structure and the dynamics of language acquisition*. Amsterdam: John Benjamins, pp. 268-287.
- Carroll, M. & C. von Stutterheim (2003). Typology and information organisation: perspective taking and language-specific effects in the construal of events. In A. Giacalone Ramat (ed.) *Typology and second language acquisition*. Berlin/New York: Mouton de Gruyter, pp. 365-402.
- Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (ed.) *Subject and topic*. New York: Academic Press, pp. 25-55.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Costa, J. (2004). *Subject positions and interfaces*. Berlin: Mouton de Gruyter.
- Daneš, F. (1970). Zur linguistischen Analyse der Textstruktur. *Folia Linguistica* 4, pp. 72-79.
- Fanselow, G. (2007). The restricted access of information structure to syntax: A minority report. In C. Féry, G. Fanselow & M. Krifka (eds.) *Interdisciplinary Studies on Information Structure* 6, pp. 205-220.
- Féry, C. (2007). Information structural notions and the fallacy of invariant correlates. In C. Féry, G. Fanselow & M. Krifka (eds.) *Interdisciplinary Studies on Information Structure* 6, pp. 161-184.
- Halliday, M. & R. Hasan (1976). *Cohesion in English*. London: Longman.
- Jackendoff, R. (2002). *Foundations of language*. Oxford: Oxford University Press.
- Jacobs, J. (1984). Funktionale Satsperspektive und Illokutionsemantik. *Linguistische Berichte* 91, pp. 25-58.
- Krifka, M. (2007). Basic notions of information structure. In C. Féry, G. Fanselow & M. Krifka (eds.) *Interdisciplinary Studies on Information Structure* 6, pp. 13-55.
- Kuno, S. (1987). *Functional syntax: Anaphora, discourse, and empathy*. Chicago: Univ. of Chicago Press.

- Lambrecht, K. (1994). *Information structure and sentence form: topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Liceras, J. (1988). Syntax and stylistics: More on the pro-drop parameter. In J. Pankhurst, M. Sharwood-Smith & P. Van Buren (eds.) *Learnability and second languages: A book of readings*. Dordrecht: Foris, pp. 71-93.
- Lozano, C. (2006). Focus and split-transitivity: the acquisition of word order alternations in non-native Spanish. *Second Language Research* 22, pp. 145-187.
- Neeleman, A. & H. van de Koot (2008). Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11 (2), pp. 137-189.
- Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole (ed) *Radical pragmatics*. New York: Academic Press, pp. 249-264.
- Prince, E. (1998). On the limits of syntax, with reference to left dislocation and topicalization. In P. Culicover & L. McNally (eds.) *Syntax and Semantics*. Vol. 39. *The limits of syntax*. New York: Academic Press, pp. 281-302.
- Reinhart, T. (1982). *Pragmatics and linguistics: A analysis of sentence topic*. Bloomington: Indiana University Linguistics Club, pp. 1-38.
- Reis, M. (1980). On justifying topological frames: 'Positional field' and the order of nonverbal constituents in German. *Documentation et Recherche en Linguistique Allemande Contemporaine* 22/23, pp. 59-85.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (ed.) *Elements of grammar: Handbook of generative syntax*. Dordrecht: Kluwer, pp. 281-337.
- Rosén, C. (2006). *Warum klingt das nicht deutsch? Probleme der Informationsstrukturierung in deutschen Texten schwedischer Schüler und Studenten*. PhD Dissertation, Stockholm: Almqvist & Wiksell.
- Teleman, U., S. Hellberg & E. Andersson (eds.) (1999). *Svenska Akademiens grammatik*. Vol. 4. *Satser och meningar*. Stockholm: NordstedtsOrdbok.
- Tsimpili, I.-M. and A. Sorace (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia & C. Zaller (eds.) *BUCLD Proceedings 30*. Somerville, MA: Cascadia Press, pp. 653-664.
- White, L. (2009). Grammatical theory: Interfaces and L2 knowledge. In W. C. Ritchie & T. K. Bhatia (eds.) *The New Handbook of Second Language Acquisition*. Leeds: Emerald.
- Zifonun, G., L. Hoffmann & B. Strecker (eds.) (1997). *Grammatik der deutschen Sprache. Schriften des Instituts für deutsche Sprache* Vol. 7.2. Berlin/New York: Walter de Gruyter.