

*Title page.*

*Title:*

**Focus types and argument asymmetries**

**A cross-linguistic study in language production**

*Authors:*

Stavros Skopeteas and Gisbert Fanselow

University of Potsdam

*Running head:*

Focus types and argument asymmetries



## **Focus types and argument asymmetries**

### **A cross-linguistic study in language production**

Stavros Skopeteas and Gisbert Fanselow

#### **Abstract**

The effects of focus on syntax differ across languages: some languages encode focus in situ, while in other languages focus induces an array of constructions that deviate from the canonical configuration, such as non-canonical orders or clefts. This article presents semi-spontaneously produced data from American English, Québec French, Hungarian, and Georgian that shows exactly that speakers of these languages select different structures in identical discourse conditions. The observed cross-linguistic differences are accounted for by means of grammatical properties of the object languages that hold independently of information structure. This account leads to the conclusion that a non-compositional mapping between information structural concepts and structural configurations is an unnecessary complication of the grammatical model.

## **1. Preliminaries<sup>1</sup>**

Previous work on information structure has identified two asymmetries with respect to the realization of focused constituents. The first asymmetry relates to the *focus type*, i.e. the type of contribution the focused constituent makes to the discourse context. Though there is a variety of functional concepts that have been used in order to establish classifications of focus (see Dik 1997, Siewierska 1991, Gussenhoven 2007, and Krifka 2007 for some detailed classifications), there is a major division between those instances of focus that simply express non-presupposed information and those that come with an additional function that operates on the relation between the focused constituent and its antecedent(s) in discourse. Following É. Kiss (1998: 262), we use the term ‘identificational focus’ for the latter variety and we assume that this type of focus involves a quantificational operation over a set of referents, in particular an operation excluding some (contrastive) or all (exhaustive) relevant alternative referents to the focused element in discourse. We use the term ‘non-identificational focus’ for the former instances of focus that do not bear any quantificational properties (also called ‘information focus’, see É. Kiss 1998). Cross-linguistically, it has been claimed that these focus types differ in their structural realization. In general, deviations from the canonical syntactic configuration are more likely to be induced by the identificational instances of focus than by the non-identificational ones. Some syntactic

models capture this asymmetry by assuming that non-canonical syntactic configurations arise through the application of some syntactic operation that is associated with identificational focus (or a subtype of it) (see É. Kiss 1998, 2009, Drubig 2003). The asymmetry of focus types is summarized in the implicative relation in (1) which should be read as follows: “If a non-canonical structure occurs with the non-identificational instances of focus, it is expected to occur with identificational instances of focus”. The predictive power of (1) is that it excludes a grammar in which non-canonical structures occur with non-identificational instances of focus while identificational instances of focus are expressed through canonical structures. We conceive the asymmetry in (1) as an observational generalization. As we are going to show in the discussion of our empirical data, this asymmetry may be derived by the interaction of contextual conditions with particular structural properties of the grammars at issue.

(1) Asymmetry of focus types

Identificational focus ← Non-identificational focus

The second asymmetry that is discussed in this article relates to the *argument hierarchy*. It has already been observed for some languages that focus on subjects obligatorily induces a non-canonical structure while focus on non-subjects only optionally does so. Evidence for subject/non-subject asymmetries has been provided for several languages including French (Lambrecht 2001), Spanish (Büring & Gutiérrez-Bravo 2001), Hausa (Hartmann & Zimmermann 2007), West Chadic languages (Zimmermann

2008), several Kwa and Gur languages (Fiedler & Schwarz 2005), Northern Sotho (Zerbian 2007), etc. This asymmetry is summarized in the implicative relation in (2) which should be read as follows: “If a non-canonical structure occurs with focus on non-subjects, it is expected to occur with focus on subjects too”. This implicative relation reflects the observation that non-canonical structures for the expression of focus occur either (a) equally for subjects and non-subjects, or (b) for subjects but not for non-subjects, or (c) for neither structural category. The argument asymmetry in (2) excludes a language type in which a non-canonical construction is used for focusing non-subjects and a canonical one for focusing subjects.

(2) Asymmetry of focused arguments

Subject ← Non-subject

Similarly to (1), we conceive the asymmetry in (2) as an observational generalization. Several explanations about the rules of grammar that account for this asymmetry have been already proposed in previous literature. A straightforward account for the asymmetry in (2) is the assumption of a constraint that bans focus on subjects of canonical sentences (see Lambrecht 2001, Van Valin 1999). A further possibility would be to assume a default association ‘subject ↔ topic’ implying that deviations from this configuration should be structurally marked (see Lambrecht 2001: 490, Zerbian 2007: 336, Hartmann & Zimmermann 2007). These accounts have in common that they directly map information structural concepts on syntactic functions. Alternatively, it is possible to derive the argument

asymmetry by general properties of the linearization or the prosodic structure, such as the phonological requirement for the rightmost prosodic constituent to be the head of a phonological phrase (see effects on argument asymmetry depending on the ranking of phonological and syntactic constraints in Büring & Gutiérrez-Bravo 2001).

This article presents comparative empirical evidence from Georgian, Hungarian, American English, and Québec French. These languages form an interesting quadruple for testing hypotheses on focus-related operations. In Georgian and Hungarian, focus may induce deviations from canonical word order, while American English and Québec French display fairly rigid word order (reorderings are constructionally and stylistically restricted). Moreover, English displays a freedom in the placement of prosodic prominence which allows for the expression of focus without any syntactic operation, while Georgian, French, and Hungarian are restrictive in this respect. These differences are outlined in section 2.

A central issue in the present volume is the question of *tertium comparationis* with respect to the cross-linguistic analysis of information structure. Descriptions of information structure in different languages not only differ with respect to their theoretical foundations but also with respect to the range of data that they consider. In order to achieve comparability of the primary data, we developed an elicitation task that establishes particular context types by means of visual stimuli and minimal verbal contributions (e.g., several questions).<sup>2</sup> The use of the same elicitation procedure in all

object languages yields a data set of semi-spontaneous expressions that is ideal for the testing of cross-linguistic hypotheses. This elicitation task is presented in section 3 and the empirical results are reported in section 4.

The theoretical question of this article is whether the cross-linguistic differences that are captured by the observational generalizations in (1) and (2) reflect: (a) non-further-decomposable differences of the individual grammars with respect to the association of information structural concepts with structural operations or (b) the interaction of universal information structural principles with structural differences of the grammars at issue. From a conceptual viewpoint, an account of the latter type has the theoretical advantage of being less stipulative, since it explains discourse-related phenomena on the basis of structural rules that independently hold. To the extent that a compositional account of this type is possible, it will give further support to the view that the correlation between information structural concepts and structural operations is not the result of a non-further decomposable ‘discourse:syntax’ association but rather the product of the interaction of discourse-related principles with the output of syntactic rules, i.e. particular linearizations and prosodic possibilities (see Wedgwood 2003, Fanselow 2006, 2007, Fanselow & Lenertová 2008, Zimmermann 2007). Nevertheless, the possibility of an account of this type is an empirical question that is discussed in section 5.

## 2. Strategies for expressing narrow focus

### 2.1 *Focus in situ*

A source of cross-linguistic variation that interacts with information structure relates to the possibility in a particular grammar to express focus *in situ*. This property probably depends on prosodic constraints (e.g., the possibility of deviating from the default prosodic structure, see Büring & Gutiérrez-Bravo 2001) that are beyond the scope of this article (see Féry & Greif 2009 for a prosodic account on the same data set). For our purposes, it is important to distinguish between languages that may express focus *in situ* and those that do not, since this possibility interacts with focus-related syntactic operations: if focus may be expressed *in situ* (through prosodic prominence), then *ex situ* focus occurs only in a subset of the instances in which a constituent bears a focus feature. This implies that focus is not a *sufficient* condition for triggering the related syntactic operation.

English is the textbook example of a language with free focus placement, i.e. any constituent may be rendered prosodically prominent *in situ* (see Gussenhoven 2007 and references therein).

Georgian is certainly restrictive in comparison to English<sup>3</sup>, but previous research has shown that there are two alternative realizations of narrow focus that do not differ in their interpretational properties (see Skopeteas &

Fanselow 2009b for detailed discussion). The one option is to express focus *in situ* and the other option to apply a movement operation (see section 2.2). Spoken French is known to have a constraint against preverbal foci (Lambrecht 2001: 492). This observation is in line with the prosodic properties of French, in particular with the fact that prosodic prominence in this language is obligatorily realized at the right edge of the Phonological Phrase and cannot be displaced from this default position in order to signal focus on non-phrase-final constituents *in situ*. Féry (2001) argues that French does not display pitch accents for the signaling of focus, a property which is traced back to the absence of lexical stress in this language. In this view, the prosodic prominence at the right edge of the Phonological Phrase is a correlate of phrasing, i.e. a boundary tone, and not a pitch accent. Crucial for our purposes is that French does not use the possibility of free pitch accent placement in order to signal that a non-phrase-final constituent is focused.

Following É. Kiss (1998: 249), *in situ* constituents in Hungarian cannot be identificationally focused. However, recent work by Szendrői (2001, 2003) shows that the asymmetry between the position immediately preceding the predicate and the postverbal domain can be traced back to properties of the prosodic structure of Hungarian utterances: movement to the preverbal position is the only possibility for a constituent to receive prosodic prominence.

## 2.2 Reordering

A further source of cross-linguistic variation with respect to the expression of information structure relates to the structural possibility of a grammar to allow for alternative linearizations of the same constituents. Two structural operations are at issue: (a) instances of movement to A-bar positions that are headed by functional projections outside the lexical domain, and (b) instances of scrambling within the lexical domain of the hierarchical clause structure.

Hungarian is a language with VSO canonical order. The occurrence of a constituent in a preverbal position is licensed by restricted contextual conditions. Two configurations involving preverbal realization of constituents have to be distinguished, see (3a) and (3b). Example (3a) could occur in a context with a subject topic (e.g., as an answer to the question ‘What did Mary do?’), while example (3b) could occur in a context that licenses narrow focus on the subject (e.g., as an answer to the question ‘Who called up Peter?’). In both cases, the subject constituent surfaces in a position that precedes the predicate. However, the preverb *fel* ‘up’ surfaces in its default position in (3a), while in (3b) it surfaces postverbally.

(3) Hungarian (É. Kiss 1998: 256)

a. *Mari fel hívta Pétert.*

Mary up called Peter.ACC

‘Mary called up Peter.’

b. *Mari hívta fel Pétert.*

Mary called up Peter.ACC

‘It was Mary that called up Peter.’

The phenomenon illustrated through (3a-b) is the basic evidence for distinguishing two preverbal positions in Hungarian. Topics are realized in a sentence-initial position which is identified by the fact that it precedes the landing site of focused constituents (see examples in É. Kiss 1998). Focused constituents undergo movement to the specifier position of another functional projection, whose head attracts the V to the effect that the latter precedes the preverb in the linear order (see É. Kiss 1998: 256). Both preverbal positions are not argument positions, i.e., they are A-bar positions above the predicate phrase. The range of contexts that induce the operation exemplified in (3b) is a matter of debate. Some accounts assume that this position is associated with a quantificational operator encoding exhaustive identification of the moved constituent (see É. Kiss 1998), while other accounts assume that this position is semantically underspecified (see Wedgwood 2003, 2007).

Georgian is a verb final language (SOV) allowing for considerable word order freedom determined by information structure (see Apridonidze, 1986: 136-143; Vogt, 1971: 222, Skopeteas & Fanselow 2009a, 2009b; Skopeteas et al. 2009). As expected for V-final languages (see Haider & Rosengren 2003), Georgian allows for word order changes of the scrambling type

which means that movement targets argument positions, which is empirically supported by the fact that the non-canonical orders establish new binding relations (see evidence and discussion in McGinnis 1999, Skopeteas & Fanselow 2009a). The interaction of scrambling with the focus set of the utterance is exemplified in (4): (4a) is a canonical SOV sentence that could be an answer to the question ‘What happened?’ (all focus) or ‘What did a/the man do?’ (VP focus) or ‘What did a/the man push?’ (object focus). (4b) illustrates a sentence in which the object is scrambled over the subject constituent. This order is contextually restricted, i.e., it could be the answer to the question ‘Who pushed the chair?’ (subject focus) (see experimental evidence as well as competence data on Georgian word order in Skopeteas & Fanselow 2009b). Speakers’ intuitions indicate that the SOV linearization in (4a) is not felicitous in subject focus contexts and that the OSV linearization in (4b) is not felicitous in object focus contexts (see Skopeteas et al. 2009, Skopeteas & Fanselow 2009b). Hence, the generalization in the Georgian data is that a preverbal constituent in narrow focus has to be realized adjacent to the verb.

(4) Georgian

a. *k’ac-i sk’am-s a-c ’v-eb-a.*

man-NOM chair-DAT PV(IO.3)-push-THM-S.3.SG

‘A/the man pushes a/the chair.’

b. *sk’am-s k’ac-i a-c ’v-eb-a.*

chair-DAT man-NOM PV(IO.3)-push-THM-S.3.SG



man-NOM PV(IO.3)-push-THM-S.3.SG chair-DAT

‘A/the man pushes a/the chair.’

English and French display a number of constructions that involve deviations from the canonical word order. However, it generally holds that reordering in these languages is restricted to particular types of constructions (e.g., the quotative inversion or the locative inversion) and is partially restricted to particular registers (e.g., French clitic constructions and the related predicate-subject order are characteristic of spoken French, see De Cat 2005: 1195). English allows for several types of reordering, including preposing, postposing, left- and right-dislocation and argument reversal (see Birner & Ward 2004). For the purposes of our article, it is relevant that object preposing may be used to express identificational focus; however, it should be noticed that this construction is generally characterized as “marked” in English, which implies that it only occurs in a very limited type of contexts/discourse situations and is associated with processing difficulty (see Breul 2007).

### *2.3 Cleft constructions*

An alternative means for expressing narrow focus is the formation of a cleft construction (see Rochemont 1986: 127ff., Lambrecht 2001 among others). The syntactic analysis of cleft constructions opens a long array of theoretical possibilities (for a summary, see Hedberg 2000: 907-912), that do not

necessarily motivate different assumptions about the information structural properties of clefts that are dealt with in this article. The crucial point for the analysis of our data is the assumption that the cleft construction in (6a) and its canonical counterpart in (6b) may be used to describe the same situation.

- (6) American English
- a. *It's a man that's pushing the car.*
  - b. *A man is pushing the car.*

English cleft constructions are used in discourse in order to realize a partition of the utterance into an asserted part, which is the clefted constituent, and a presupposed part, which surfaces as a relative clause. It is generally assumed that the clefted constituent is identificationally focused (see É. Kiss 1998: 268, Lambrecht 2001: 497, Rochemont 1986: 133). This property can be implemented in monoclausal accounts of cleft constructions quite straightforwardly, by assuming that the landing site of movement is associated with a particular operator (e.g., the operator [+ exhaustive] in É. Kiss 1998: 268). Alternatively, the focus properties may be accounted for in terms of independent interpretative principles (see 'Cleft Focus Principle' in Rochemont 1986: 133) which can apply to any syntactic account on cleft sentences.

Drubig (2003) assumes that the interpretative properties of clefts are directly derived from the syntactic configuration. Cleft sentences instantiate movement to a specifier position within the complementizer layer of the clause (CP) and this operation is associated with a contrastive reading (see

Drubig 2003: 14). An apparent problem for this assumption is that constructions that have (at least) the superficial properties of clefts do not have identical interpretational properties across languages. For instance, the corresponding French construction in (7), though superficially identical to the English example (6a) does not display the same focus possibilities. Lambrecht (2001) argues that this construction occurs whenever the subject is part of the focus domain (including cases of narrow and broad focus), a hypothesis that is experimentally confirmed for Québec French in Thériault et al. (2008). In this view, the example in (7) could be an answer to the question ‘Who is pushing the car?’ (subject focus) or ‘What happens?’ (all focus, i.e. subject is part of the broad focus domain). The crucial theoretical question is where the interpretative difference between superficially identical constructions in different languages comes from (see further discussion in section 5).

(7) Québec French

*C' est un homme qui pousse*

it be:3.SG INDEF:M.SG man who push:3.SG

*l' auto.*

DEF.M.SG car

‘It is a man that pushes the car.’

The two further languages in our sample, namely Hungarian and Georgian, also have the structural possibility to form cleft constructions (the corresponding constructions in these languages are reversed pseudo-clefts).

However, these constructions occur only rarely in discourse and native speakers' intuitions suggest that they are restricted to specific registers ("written styles").

### **3. Method**

The aim of the elicitation task that is presented in this section is to create a semi-naturalistic data set that allows us to observe the effects of the asymmetries presented in section 1. This elicitation task is part of the *Questionnaire on Information Structure* (see section 1, footnote 2). The experimental procedure is based on the elicitation of spontaneous answers to several question types. The speaker is presented four pictures and is instructed to look at the presented scenes. When (s)he is ready, the pictures are taken away and the instructor asks four questions concerning the perceived stimuli. The speaker is instructed to avoid elliptical answers such as "yes", "no", "the man", etc., and to give a syntactically complete answer to the question instead.

The examined factors correspond to the asymmetries introduced in section 1. The factor 'focused argument' is intended to provide evidence for the asymmetry between focus on subjects and focus on non-subjects, see (2). The factor 'focus type' is intended to provide evidence for the asymmetry between non-identificational and identificational foci, see (1). The

permutation of the levels of both factors results in four experimental conditions that are listed and exemplified in (8a-d). The non-identificational conditions involve *wh*-questions that induce narrow focus on the subject or object constituent, see (8a-b). The underlying assumption is that *wh*-questions do not trigger an answer that involves an explicit expression of exhaustive identification. The possible exhaustive interpretation of the answer in this context is independent of its form, i.e., it is available also with answers in the canonical order. This interpretation is the result of a pragmatic inference that is motivated by the fact that the *wh*-question is interpreted as a request to assert the exact subset of referents for which the proposition holds (see Groenendijk & Stokhof 1984) and the assumption that the utterer of the answer is cooperative, i.e., (s)he observes the request in the conversational context. The questions in the identificational conditions induce an answer that involves contrast to either the subject or the object constituent, see (8c-d).

(8) Conditions

Stimulus: ‘in front of a well, a man is pushing a car’

- a. Condition N/SBJ: non-identificational, subject  
{In front of the well, who is pushing the car?}
- b. Condition N/OBJ: non-identificational, object  
{In front of the well, what is the man pushing?}
- c. Condition I/SBJ: identificational, subject  
{In front of the well, is a woman pushing a car?}

- d. Condition I/OBJ: identificational, object

{In front of the well, is the man pushing a bicycle?}

Each participant of the experiment was presented four picture sheets, containing four pictures each, hence each participant produced a total of  $4 \times 4 = 16$  answers. Half of these questions correspond to the conditions in (8), which means that we elicited two answers for each question type per speaker.<sup>4</sup> The tasks were pseudo-randomized and part of a longer elicitation session that contained several tasks of the *Questionnaire on Information Structure*.

#### **4. Results**

The effects of focus on the clause structure may be tested in the subset of answers that (i) realize the intended contextual conditions and (ii) involve a lexically realized verb. Answers that do not meet these requirements were coded as ‘non-valid’ and are discarded in the further analysis (which means that they are natural answers in the examined discourse condition, but irrelevant for the hypotheses at issue). (9a) illustrates an answer in the English data set that does not meet requirement (i) and (9b) an answer in the Georgian data set that does not meet requirement (ii). The observations made in the following sections are based on the remaining answers that were decoded as ‘valid’.

- (9) a. {In the scene with cloudy sky, who is looking at the girl?}  
*Who is looking at the g...? The man is looking at the girl?* (Condition N/SBJ)
- b. {In front of the well, who is pushing a/the man?}  
*bič'-i.*  
boy-NOM  
'A/the boy.' (Condition N/SBJ)

#### 4.1 Georgian<sup>5</sup>

In the set of valid data, we encountered two types of realization of the focused constituent. The first type consists of sentences in which the focused constituent (either subject or object) is placed in the immediately preverbal position, which is the case in the orders  $SO_FV$ ,  $O_FVS$ ,  $OS_FV$ ,  $S_FVO$ ,  $O_FV$ , and  $S_FV$  (see Table 1). The crucial observation is that while the  $(X)Y_FV$  pattern occurs in several configurations, the  $X_FYV$  pattern is not attested at all. This contrast provides evidence for the generalization that a preverbal constituent in narrow focus has to be realized adjacent to the verb, see 2.2. The following examples illustrate two deviations from the canonical SOV order (see Skopeteas & Fanselow 2009b for further examples and discussion of this data set): the focused subject in (10a) is realized adjacent

to the verb in an OS<sub>F</sub>V order; in (10b), the focused object is left adjacent to the verb, while the given argument is realized postverbally.

- (10) a. OS<sub>F</sub>V
- {In the scene with the blue sky, is a/the man hitting a/the man?}
- ara, k'ats-s kal-i*
- no man-DAT woman-NOM
- u-rt'q'-am-s.*
- PV(IO.3)-hit-THM-S.3.SG
- 'No, a/the woman is hitting a/the man.' (Condition I/SBJ)
- b. O<sub>F</sub>VS
- {In the scene in the room, what is a/the man hitting?}
- sk'am-s u-rt'q'-am-s igi.*
- chair-DAT PV(IO.3)-hit-THM-S.3.SG that:NOM
- 'He is hitting a/the chair.' (Condition N/OBJ)

The second option of realization of the focused constituent in Georgian is postverbal, as exemplified in (11a) for SVO<sub>F</sub> and (11b) for OVS<sub>F</sub>. Following our account in Skopeteas & Fanselow (2009b), these sentences involve optional V-movement to a higher position in the hierarchical clause structure, see discussion in section 2.2. Hence, the focused constituent in both examples is realized *in situ*.

- (11) a. SVO<sub>F</sub>

{In the scene in front of the fence, what is a/the girl hitting?}

*gogo u-rt'q'-am-s mankana-s.*

girl(NOM) PV(IO.3)-hit-THM-S.3.SG car-DAT

'A/the girl is hitting a/the car.' (Condition N/OBJ)

b. OVS<sub>F</sub>

{In the scene with the blue sky, who is hitting a/the man?}

*k'ac-s u-rt'q'-am-s kal-i.*

man-DAT PV(IO.3)-hit-THM-S.3.SG woman-NOM

'A/the woman is hitting a/the man.' (Condition N/OBJ)

[@@ Insert Table 1 here](#)

The impact of the contextual conditions on the choice among preverbal/postverbal focus is reflected in the means presented in Figure 1 (calculated on the basis of proportions of preverbal focus per speaker). A repeated-measures analysis of variance on the proportions obtained by each speaker separately revealed a significant main effect of 'focused argument' ( $F_{1,14} = 8.44, p < .01$ ) and of 'focus type' ( $F_{1,14} = 5.05, p < .04$ ) and no significant effect of the interaction between the two factors. The two main effects indicate that both factors have an impact on the occurrence of orders in which the focused constituent is placed in the immediately preverbal

position in Georgian, i.e. that focused subjects are more likely to occur in this position than focused objects, and identificational focus is more likely to occur in this position than non-identificational focus. The absence of a significant interaction indicates that the impact of these factors is independent from one another.

@@ Insert Figure 1 here

#### 4.2 *Hungarian*<sup>6</sup>

We have seen in section 2.2 that the surface placement of Hungarian preverbs provides evidence for the distinction between a sentence initial position and an immediately preverbal position. The examples in (12) illustrate the word orders in our data set (SVO, SOV, and OVS) in which the verb precedes the preverb, indicating thus that the preverbal constituent occupies the type of preverbal position in Hungarian that invokes V-attraction (compare with the preverb-verb order in (14)).

- (12) a. S<sub>F</sub> Vp O  
{Is a man hitting the man?}  
*Nem, egy nõ üti meg*  
no INDEF woman hit:3.SG PRF  
*a férfit.*

DEF man:ACC

‘No, a woman is hitting the man.’ (Condition I/SBJ)

b. S O<sub>F</sub> Vp

{What is the man kicking?}

*A férfi a széket rúgja meg.*

DEF man DEF chair:ACC kick:3.SG PRF

‘The man is kicking the chair.’ (Condition N/OBJ)

c. O<sub>F</sub> Vp S

{Whom is the man kicking?}

*Egy másik férfit rúg meg*

INDEF other man:ACC kick:3.SG PRF

*a férfi.*

DEF man

‘The man is kicking another man.’ (Condition N/OBJ)

In a further subset of our data, the speakers selected verbs without preverbs (see (13)). In these utterances, the only evidence for the properties of the position at issue is the adjacency to the verb. The distribution of these sentences in Table 2 shows that SVO sentences only occur with subject focus, while SOV/OVS sentences only occur with object focus.

(13) a. S<sub>F</sub> V O

{Who is carrying the pot?}

*Egy férfi cipeli a cserepet.*

INDEF man carry:3.SG DEF pot-ACC

‘A man is carrying the pot.’ (Condition N/SBJ)

b. O<sub>F</sub> V S

{Whom is the man carrying?}

*Egy nőt cipel a férfi.*

INDEF woman-ACC carry:3.SG DEF man

‘The man is carrying a woman.’ (Condition N/OBJ)

Example (14) is the only utterance in our data set, in which the focused constituent (object) is realized *in situ*. The preverbal realization of the preverb adds evidence that the given subject in the left periphery is not in the position that invokes V-attraction. Following discourse-configurational accounts on Hungarian syntax, the postverbal argument may only bear new information focus, which means that the answer in (14) is not contextually congruent, since the context involves correction (see É. Kiss 1998). However, example (14) displays a heavy object constituent, indicating that movement to the position that hosts focused constituents in Hungarian interacts with non-pragmatic preferences on the linearization (such as the preference for heavy constituents to be realized late in the utterance, that is known to influence Hungarian word order, see É. Kiss 2008: 445-447).

(14) Identificational focus *in situ*

{Is the woman hitting a flower?}

*Nem, a nő ki-tépi az utolsó fát*

no DEF woman out-pull:3.SG DEF last tree-ACC

*a környéken.*

DEF neighborhood-SUP

‘No, the woman is pulling out the last tree in the neighborhood.’ (Condition I/OBJ)

**@@ Insert Table 2 here**

The Hungarian data set reveals a categorical pattern as shown in Figure 2. The focused constituent is realized immediately in front of the verb and this holds for both focus types and both focused arguments examined in this elicitation task. Whenever a preverb is available, then this preverb appears postverbally which supports the view that the constituent that occurs left adjacent to the verb occupies the specifier position of a functional projection whose head attracts the verb. The only exception to this pattern is a single example in the condition of identificationally focused objects. However, we argued that there is no reason to assume that this difference depends on the examined condition, since the utterance at issue contains a heavy object constituent that is probably realized in situ for reasons that do not relate to information structure.

**@@ Insert Figure 2 here**

### 4.3 *American English*<sup>7</sup>

In the American English data, we encounter three types of sentences: canonical SVO sentences as exemplified in (15a), (b) *it*-clefts (see (15b)), and (c) presentational constructions (see (15c)). The distribution of these answer types in the conditions of the elicitation task is presented in Table 3.

- (15) a. Canonical sentence  
{Who is carrying the pot?}  
*Some guy's carrying the pot.* (Condition N/SBJ)
- b. *it*-cleft  
{Is a woman pushing the car?}  
*No, it's a man that's pushing the car.* (Condition I/SBJ)
- c. Presentational construction  
{In front of the well, who is pushing the man?}  
*There is a dark skinned man pushing a white skinned man.* (Condition N/SBJ)

The presentational constructions in (15c) are not bi-clausal, since the predicate is not expressed through a relative clause. These constructions may occur in two different discourse conditions in English: either asthetic sentences, hence in an all-new context, or involving non-exhaustive subject focus (see Lambrecht 2001: 505-507). The distribution of this construction in the experimental conditions (see Table 3) shows that they only occur if the subject constituent is in focus. This fact suggests that presentational

constructions are induced by subject focus in our data, which is in line with the view that these constructions are a way to place information that has to be stressed in the position that is assigned stress in neutral prosodic structures, i.e. phrase finally (cf. the pragmatic account in Birner and Ward 2004: 163).

@@ Insert Table 3 here

Figure 3 presents the percentages of *it*-clefts in the data set and shows that the only context in which this type of cleft occurs in our data set is the condition of identificational focus on subjects.

@@ Insert Figure 3 here

#### 4.4 Québec French<sup>8</sup>

The answer types in the Québec French data set are presented in (16). Next to the canonical type of sentences in (16a), we encountered two types of cleft construction, those that are introduced by the identificational predicate *c'est* (see (16b)) and those that are introduced by the presentational predicate *y a* (see (16c)).

- (16) a. Canonical sentence

{In front of the bridge, who is carrying the pot?}

*Un homme transporte*

INDEF.M.SG man transport-3.SG

*le pot.*

DEF.M.SG pot

‘A man transports the pot.’ (Condition N/SBJ)

b. *C’est* cleft construction

{In front of the bridge, who is carrying the pot?}

*C’ est un homme qui*

it be:3.SG INDEF.M.SG man who

*transporte le pot.*

transport-3.SG DEF.M.SG pot

‘It is a man that transports the pot.’ (Condition N/SBJ)

c. *Y a* cleft construction

{In front of the bridge, who is carrying the pot?}

*Y a un homme qui*

there have:3.SG INDEF.M.SG man who

*transporte le pot.*

transport:3.SG DEF.M.SG pot

‘There is a man that transports the pot.’ (Condition

N/SBJ)

The distribution of these answer types in the examined contextual conditions is presented in Table 4. The contrast between the two types of

matrix predicate (*c'est* vs. *y a*) encodes the distinction between identificational and existential clauses in the language. In a compositional view, the occurrence of these types of predicate in what surfaces as a matrix clause in cleft constructions is expected to correlate with the contrast between identificational and presentational clefts (the former occurring in narrow focus and the latter in broad focus utterances). However, the occurrence of *y a* clefts in narrow focus constructions in Table 4 suggests that the semantic properties of the matrix predicates do not have a compositional contribution to the semantics of cleft constructions of Québec French. Both types of predicate are merely alternative lexicalizations for a particular syntactic configuration (of the cleft type).

All cleft constructions in the data set involve a clefted subject constituent; constructions with clefted objects do not occur at all. This also holds for the three cleft constructions that are encountered in object focus conditions in Table 4, exemplified in (17). The question is what triggers the marginal occurrence of clefted subjects in the context of object focus questions. Recall from section 2.3, that French clefts may occur whenever the subject constituent is part of the focus domain, including cases of narrow and broad focus. Since narrow focus on the subject does not apply in this context, we assume that the three subject cleft constructions in the context of object questions represent the (marginal) case that speakers give an out-of-the-blue description ignoring the content of the question. Under this reading, the answer in (17) is not congruent to the question, though it is an informative

contribution to the discourse (see Thériault et al. 2008 for further discussion).

(17) {In the scene with the blue sky, whom does the woman hit?}

*Y a une femme qui frappe*  
there have:3.SG INDEF.F.SG woman who hit-3.SG  
*un homme.*  
INDEF.M.SG man

‘There is a woman that hits a man.’ (Condition N/OBJ)

[@@ Insert Table 4 here](#)

Figure 4 presents the proportions of the data in which the respective focused constituent is clefted. The data pattern is different from the English one in Figure 3. First, clefting the focused constituent frequently occurs in both conditions of subject focus and only in these (recall that the three cleft constructions in the object-focus conditions involve clefted subjects, see example (17)). Second, the proportions of cleft constructions in these conditions are higher than the corresponding proportion in English (see Figure 3). A repeated measures analysis of variance at an alpha level of .05 revealed a significant main effect of the factor ‘focused argument’ ( $F_{1,9} = 34.15, p < .001$ ), no significant effect of ‘focus type’ nor of the interaction between the two factors. According to these findings, there is no evidence that the examined focus types have a distinct impact on the selection of cleft

constructions in Québec French. The asymmetry between subject and object focus has a significant impact though, such that cleft constructions are more likely to occur in the former discourse condition than in the latter.

@@ Insert Figure 4 here

## 5. Discussion

### 5.1 Summary of empirical findings

In sum, the elicitation task revealed the following empirical generalizations:

- a. *Georgian*: Narrow focus is optionally expressed through the immediately preverbal position or otherwise *in situ*; the proportions of focus in the preverbal position reveal a significant effect of argument asymmetry and a significant effect of the asymmetry of focus type.
- b. *Hungarian*: Narrow focus is always expressed *ex situ* (a single counterexample is accounted for through the influence of heaviness constraints).
- c. *American English*: Identificational focus on subjects induces a low proportion of cleft constructions.
- d. *Québec French*: All types of narrow focus on subjects induce high proportions of cleft constructions.

Apart from Hungarian, all languages display a subject/object asymmetry, such that subject focus induces a non-canonical structure more frequently. In Georgian, the pattern is probabilistic, while in American English and Québec French the pattern is categorical (only clefted subjects).<sup>9</sup> Georgian and American English reveal an asymmetry depending on focus type: identificational focus induces a greater proportion of preverbal focus in Georgian and only identificational focus induces clefts in English.

## *5.2 Interaction with grammatical possibilities*

The aim of this section is to account for these empirical differences on the basis of the grammatical background that is introduced in section 2. We assume that native speakers select a structural possibility from a set of alternative options for the expression of narrow focus that are determined by the grammar. The relevant sets for the languages at issue are given in (18). Following the grammatical information of section 2.1, Georgian and American English have the possibility to express narrow focus *in situ* through prosodic properties that apply independently of syntax. This possibility does not hold for Hungarian. In French, prosodic prominence falls by default to the sentence-final constituent, see section 2.1, which implies that objects may be focused *in situ*, while subjects do not. As introduced in 2.2, the grammar of Georgian and of Hungarian allow operations of simple reordering that may be applied in the examined

context, while the reordering options of English and French are constructionally and/or stylistically restricted and do not apply to the context at issue. Finally, all languages have the possibility to form constructions that are (at least) superficially bi-clausal.

(18) Sets of structural alternatives per discourse condition

Georgian:       $\text{sbj|obj}_{\text{FOC}} \rightarrow \{\textit{in situ}, \text{reordering}, \text{clefting}\}$

Hungarian:     $\text{sbj|obj}_{\text{FOC}} \rightarrow \{\text{reordering}, \text{clefting}\}$

Am. English:    $\text{sbj|obj}_{\text{FOC}} \rightarrow \{\textit{in situ}, \text{clefting}\}$

Q. French:      $\text{sbj}_{\text{FOC}} \rightarrow \{\text{clefting}\}$

$\text{obj}_{\text{FOC}} \rightarrow \{\textit{in situ}, \text{clefting}\}$

The grammatical properties presented in section 2 and summarized in (18) already explain a part of the obtained data patterns. They explain why we did not get any instances of *in situ* focus in Hungarian (apart from a single example with a heavy constituent), and why we obtained a subject/object asymmetry in Québec French. The asymmetry in the Québec French data refers to the difference between the proportion of clefts in the object focus condition and the corresponding proportion in the subject focus condition, which is statistically reflected on the significant main effect of the factor ‘focused argument’, see section 4.4. Apart from this difference, we obtained a substantial amount of canonical SVO sentences in all conditions. If the occurrence of a construction in an experimental condition is taken as evidence that this construction encodes the corresponding contextual

configuration, then this part of the data provides evidence that focus on subjects may be realized *in situ*, hence counterevidence to the expectation for  $\text{sbj}_{\text{FOC}}$  in (18) that corresponds to Lambrecht's constraint against preverbal foci in spoken French (see 2.1). This interpretation results in a strong claim based on the residual of the empirically attested differences. Crucially, SVO sentences are the canonical configuration in French, hence it cannot be excluded that these utterances are informative reactions to the subject question without an overt expression of subject focus, i.e., without a focus-background articulation (for a prosodic analysis of these utterances, see Féry & Greif 2009). From the empirical viewpoint, the interpretable part of the dataset refers to the obtained differences, that provide evidence that the factor 'focused argument' has a significant effect on the choice among a canonical and a cleft construction, which is in line with Lambrecht's constraint on preverbal foci and the prediction on French  $\text{sbj}_{\text{FOC}}$  in (18).

### 5.3 *Minimality condition*

The alternative structures in (18), namely *in situ*, reordering, and clefting, differ in structural complexity. In particular, *in situ* focus does not involve any syntactic operation, hence it qualifies as the least complex structure. Since a cleft construction contains additional structural material, we may plausibly assume that clefting involves a higher degree of structural

complexity than simple reordering. These considerations lead to a markedness scale that is presented in (19).

- (19) Scale of structural complexity  
*in situ* < reordering < clefting

Some properties of the data pattern are straightforwardly explained if we assume that speakers' choices are guided by economy. The concept of economy that applies to the type of data from language production reflects the same fundamental assumptions with the concept of economy in derivational syntax (see Chomsky 1992: 47f.). In the production data presented here, economy determines the speaker's choice among existing structural alternatives for the expression of the same propositional content, reflecting the 'least effort' principle which has been shown to account for several properties of language processing (see Bornkessel & Schlesewsky 2007 and references therein); it has already been observed that the asymmetry between optimal and suboptimal structures has an even stronger effect in production data since optimal candidates always win the competition to their alternatives in discourse (see Featherston 2005). In this spirit, we formulate the minimality condition for the production data as follows.

- (20) Minimality condition (for language production)

If two structures  $s_1$  and  $s_2$ , such that  $s_1 < s_2$  in structural complexity, may be used for the same information structural configuration, the speaker selects  $s_1$ .

The minimality condition accounts for a further subset of the empirically attested differences. It explains why French speakers did not use cleft constructions in the object focus conditions, as well as why cleft constructions are not attested at all in the Georgian and Hungarian data sets (though they are possible structural configurations in the grammar). I.e., with the assumptions made so far, we may completely account for the data pattern in Hungarian and Québec French, but not yet for the subject/object asymmetry and the identificational/non-identificational asymmetry in the Georgian and American English data.

#### *5.4 Asymmetry of focused arguments*

Based on the grammatical background on Georgian in section 2.2, the subject/object asymmetry in this language may be accounted for as an interaction of the canonical word order properties with the obligatory V-attraction. *In situ* focus on object constituents may be realized in the canonical SOV order, while *in situ* focus on the subject is not possible in the SOV order and may appear only in sentences in which the V is raised in a higher clausal position.

In English, focus on the object is realized through prosodic prominence on the sentence-final constituent which corresponds to the default prosodic structure in this language (see Gussenhoven 2007), while *in situ* focus on the subject has to be realized by a non-canonical prosodic structure. This

difference implies that the choice of an alternative strategy will be more likely whenever non-sentence final constituents are in focus. The fact that English provides prosodic means to signal *in situ* focus in contrast to French is reflected on the difference between the proportion of cleft sentences in American English (21.4% in identificational subject focus contexts) and the proportion of cleft sentences in Québec French (54.7% in non-identificational subject focus and 74% in identificational subject focus). However, since the non-sentence final constituents in our data set are subjects, the evidence for an asymmetry depending on the ‘focused argument’ can be due to further structural differences as well. It is known that extraction out of relative clauses is subject to locality constraints, such that extraction of lower constituents (in our case, objects) is less likely than extraction of higher constituents (in our case, subjects). The four examined discourse conditions do not allow us to disentangle between these confounded factors. However, the crucial point for our considerations is that the obtained subject/object asymmetry in American English may be accounted for on the basis of structural differences and does not imply a non-compositional constraint on the mapping between the information structural concept of focus and the syntactic status of the focused constituent.

### *5.5 Asymmetry of focus types*

English (in subject focus) and Georgian reveal an asymmetry depending on focus type: (a) identificationally focused constituents are realized more frequently in the preverbal position in Georgian and (b) only identificationally focused subjects invoke cleft sentences in English. This data shows that information structural categories like the distinction between identificational/non-identificational focus have an impact on syntax. However, models based on the idea of biunique associations between information structural concepts and syntactic operations cannot explain our data in a straightforward manner. Two properties of our data count against the assumption of biunique association: (a) the effect of the identificational focus on the preverbal placement of a constituent in Georgian is weaker than the corresponding effect in the Hungarian data set, which suggests that the operation we observe in Georgian is an optional choice; (b) the effect of the identificational focus on clefting in English is weaker than the corresponding effect in the French data set, which suggests that clefting identificationally focused subjects is optional in English; (c) the selection of the structures at issue is also sensitive to further asymmetries in the language (such as the subject/object asymmetry) that relate to structural properties, as shown in section 5.4. These properties cannot be straightforwardly accounted for through the assumption of a biunique association between information structural distinctions and syntactic configurations.



The interpretational properties show that both the canonical and the non-canonical options of expressing focus in English and Georgian allow for the inference of the exhaustive identification. Hence, English and Georgian differ from Hungarian, in which postverbal constituents do not exhibit exhaustive readings (see É. Kiss 1998, 2009). The interpretational evidence supports the view that identificational focus is not a sufficient condition for the licensing of the non-canonical structures in these languages, as already suggested by the data pattern of our elicitation task.

For these reasons, we assume that the empirically attested asymmetry of focus types in Georgian and English does not reflect the non-compositional association between the feature [+ identificational] and particular syntactic operations, but a contextual asymmetry resulting in a probabilistic correlation with certain types of answers. *Wh*- questions (conditions N/SBJ, N/OBJ) introduce a variable and a presupposition. Answers that only assert the referent that instantiates the variable are highly expected, i.e., their information structure is fully predictable by the context, even if it is not signaled by grammatical means. The conditions I/SBJ and I/OBJ, on the other hand, involve rejection of a part of the presuppositions of the speaker, hence involving a focus feature that is not predictable by the question. By means of this asymmetry we may reasonably assume that the latter context is more likely than the former to induce a structure that articulates the focus domain at issue. The empirical confirmation of this prediction is the significant main effect of ‘focus type’ in Georgian and American English. During the

production process, this asymmetry interacts with markedness constraints resulting in a data pattern that contains a larger proportion of violations of the minimality condition in the identificational contexts. The empirical proof of this expectation is the significant interaction effect between ‘focus type’ and ‘focused argument’ in American English.

## **6. Conclusions**

The semi-spontaneous data presented in this article shows that the asymmetry of focus types and the asymmetry of focused arguments have cross-linguistically different effects on the choice of syntactic structure. We accounted for the obtained differences by means of grammatical differences between the languages at issue, notably the possibility of expressing narrow focus *in situ* and the availability of operations that allow for the expression of focus through simple manipulation of the linear order. By assuming a minimality condition in language production we were able to predict the preference for structurally less complex operations whenever they compete with more complex alternatives in particular contexts. By assuming a difference between identificational and non-identificational contexts, we predicted that the former are more likely than the latter to license violations of minimality in the speakers’ choices.

In sum, we were able to explain the properties of the behavioral data set on the basis of structural differences between the observed languages and without recourse to the assumption of associations between certain information structural concepts and particular syntactic operations. The obtained data provides evidence against a cross-linguistic 1:1 mapping between types of focus and structural operations. Hence, while French clefts occur whenever the subject is part of whatever focus domain, English clefts only occur in contexts that license an identificationally focused subject. The empirical data shows that English clefts occur in different contextual conditions than focus movement in Hungarian, which is counterevidence to the assumption that both structures are licensed by the same feature of exhaustive identification (see É. Kiss 1998). The difference in our data is in line with the conclusion of Wedgwood et al. (2006) that the range of interpretations and corpus occurrences of focus movement in Hungarian has a significantly underspecified semantics in comparison to English clefting. In our view, this difference is accounted for by the fact that English has an *in situ* alternative for signaling narrow focus, while Hungarian does not, and furthermore by the fact that the choice between an *in situ* alternative and a cleft construction interacts with structural factors.

This argumentation in this article advocates the line of thought that a substantial portion of the attested cross-linguistic differences on the effects of information structure on syntax is explained if we take into account the structural possibilities of the grammars at issue and their interaction with

communicative intentions in discourse. To the extent that these effects are predictable through structural generalizations, a non-compositional mapping between information structural concepts and structural operations leads to an unnecessary contamination of the constituent structure with pragmatic concepts.

## **Notes**

<sup>1</sup> The present article evolved within the project D2 *Typology of Information Structure*, which is part of the SFB 632 *Information Structure* at the University of Potsdam/Humboldt University Berlin (financed by the German Research Foundation). We would like to thank Carsten Breul, Caroline Féry, Edward Goebbel, Sam Hellmuth, Manfred Krifka, and Malte Zimmermann for their comments on the interpretation of the experimental data and on previous versions of this article. Special thanks are due to Rusudan Asatiani, Alain Thériault, Elizabeth Medvedovsky, and Krisztián Tronka, who contributed to the data collection and the analysis of the data sets of the individual languages. This article was presented at the conference *Contrastive Information Structure Analysis* (Wuppertal, 18 March 2008).

<sup>2</sup> The task presented in this paper is part of a longer elicitation agenda, namely the *Questionnaire on Information Structure* (QUIS), which is the collaborative product of the project *Typology of Information Structure* at the University of Potsdam/Humboldt University Berlin (see Skopeteas et al. 2006).

<sup>3</sup> The analysis of the prosodic properties of Georgian is a matter of ongoing research by Caroline Féry in association with Rusudan Asatiani and Stavros Skopeteas that we do not anticipate in this paper.

<sup>4</sup> The data presented in this paper is part of a larger data set that contains two further question types (selection and confirmation) and has been carried out in 15 languages. A full account of the obtained data is under preparation (see Skopeteas and Fanselow 2008c for syntax and Féry and Greif 2008 for prosody).

<sup>5</sup> A first dataset with 4 speakers was recorded and transcribed by Rusudan Asatiani (January-June 2005). A second dataset containing 16 further speakers was collected by S. Skopeteas and transcribed by Sh. Bartaia and N. Tsereteli (September 2005). All participants are native speakers of Georgian and residents of Tbilisi (11 women, 9 men, age range: 18-26, average: 21.9).

<sup>6</sup> The Hungarian data was collected and transcribed by Krisztián Tronka (Piliscsaba, Hungary, 2006-2007). Four native speakers participated to the experiments, all residents of Piliscsaba and students.

<sup>7</sup> The data was collected and transcribed by Elizabeth Medvedovsky (Chicago, December 2005). 20 native speakers (age range 20-26), all inhabitants of Chicago participated in the elicitation task.

<sup>8</sup> The data was collected and transcribed by Alain Thériault in Montreal (August-December 2007). 10 speakers (4 men, 6 women; age range: 25–49; average: 34.6) participated in the experiment, all residents of Montreal, native speakers of Québec French and bilingual in English. Each speaker has been presented the entire set of questions (hence gave 8 tokens for each experimental condition), which resulted in a larger data set (total: 320 answers).

<sup>9</sup> Recall that also the clefts that were encountered in the object-focus contexts involved a clefted subject constituent, see section 4.4.

<sup>10</sup> Note that these examples involve case inversion which is licensed by the perfect tense, i.e., the agent constituent bears dative case and the patient constituent nominative case (see Harris 1981).

## **References**

- Anderson, S. 1984. On representations in morphology: case, agreement, and inversion in Georgian. *Natural Language and Linguistic Theory* 2: 157-218.
- Apridonidze, Sh. 1986. *sit'q'vatganlageba axal kartulši* [word order in Modern Georgian]. Mecniereba, Tbilisi.
- Birner, B. & Ward, G. 2004. Information Structure and Non-canonical Syntax. In *The Handbook of Pragmatics*, L.R. Horn & G. Ward (eds), 153-174. Oxford: Blackwell.
- Bornkessel, I. & Schlesewsky, M. 2006. The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113(4): 787-821.
- Breul, C. 2007. Focus structure, movement to spec-Foc and syntactic processing. In Schwabe and Winkler (eds.), 255-274.

- Büring, D. & Gutiérrez-Bravo, R. 2001. Focus-related constituent order variation without the NSR: A prosody-based crosslinguistic analysis. *Syntax and semantics at Santa Cruz* 3: 41-58.
- Chomsky, N. 1992. *A Minimalist Program for Linguistic Theory*. Cambridge MA: MIT.
- De Cat, C. 2005. French subject clitics are not agreement markers. *Lingua* 115: 1195-1219.
- Dik, S. 1997. *The Theory of Functional Grammar, Part 1: The Structure of the Clause* (ed. by Kees Hengeveld). Berlin: Mouton De Gruyter.
- Drubig, H.B. 2003. Toward a typology of focus and focus constructions. *Linguistics* 44(1): 1-50.
- É. Kiss, K. 1998. Identificational vs. information focus. *Language* 74(2): 245-273.
- É. Kiss, K. 2008. Free word order, (Non)configurationality, and Phases. *Linguistic Inquiry* 39(3): 441-475.
- É. Kiss, K. 2009. Structural focus and exhaustivity. To appear in *Information Structure from different perspectives*, M. Zimmerman & C. Féry (eds). Oxford: Oxford University Press.
- Fanselow, G. 2006. On pure syntax (uncontaminated by information structure). In *Form, Structure and Grammar*, P. Brandt, E. Fuss (eds), 137-157. Akademie Verlag, Berlin.
- Fanselow, G. 2007. The restricted access of information structure to syntax: A minority report. In *The Notions of Information Structure*, C. Féry, G.

- Fanselow, M. Krifka (eds), 205-220. Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS), 6. Potsdam: Universitätsverlag.
- Fanselow, G. & Lenertová D. 2008. Left peripheral focus. Mismatches between syntax and information structure (submitted).
- Featherston, S. 2005. The decathlon model: Design features for an empirical syntax. In *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, M. Reis & S. Kepsler (eds), 188-208. Mouton de Gruyter, Berlin.
- Féry, C. 2001. Focus and phrasing in French. In *Audiatur Vox Sapientiae: A Festschrift for Arnim von Stechow*, C. Féry & W. Sternefeld (eds), 153-181. Berlin: Akademie.
- Féry, C. & Greif, M. 2009. Information structure and prosody. To appear in *The Expression of Information Structure: The Interaction of Syntax and Phonology in Cross-Linguistic Perspective*, S. Skopeteas, S. Hellmuth, G. Fanselow & C. Féry (eds). Berlin: De Gruyter.
- Fiedler, I. & Schwarz, A. 2005. Out-of-focus encoding in Gur and Kwa. *ISIS (Interdisciplinary Studies on Information Structure)* 3: 111-142.
- Groenendijk, J. & Stokhof M. 1984. Studies on the Semantics of Questions and the Pragmatics of Answers. Ph. D. thesis, Universiteit van Amsterdam.

- Gussenhoven, C. 2007. Types of focus in English. In *Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation*, C. Lee, M. Gordon, D. Büring (eds), 83-100. Heidelberg etc.: Springer.
- Haider, H. & Rosengren, I. 2003. Scrambling: Nontriggered chain formation in OV languages. *Journal of Germanic Linguistics* 15(3), 203-267.
- Harris, A.C. 1981. *Georgian syntax: A study in Relational Grammar*. Cambridge: Cambridge University Press.
- Hartmann, K. & Zimmermann M. 2007. In place – out of place? Focus strategies in Hausa. In Schwabe and Winkler (eds.), 365-403.
- Hedberg, N. 2000. The referential status of clefts. *Language* 76(4), 891-920.
- Krifka, M. 2007. Basic notions of information structure. In *The Notions of Information Structure* (Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure 6), C. Féry, G. Fanselow, M. Krifka (eds), 13-55. Potsdam: Universitätsverlag.
- Lambrecht, K. 2001. A framework for the analysis of cleft constructions, *Linguistics* 39(3): 463-516.
- McGinnis, M. 1999. A-scrambling exists! In *Proceedings of the 23th Annual Penn Linguistics Colloquium. Penn Working Papers in Linguistics*, M. Minnick and N. Han (eds). Department of Linguistics, University of Pennsylvania.
- Rochemont, M.S. 1986. *Focus in Generative Grammar*. Amsterdam: Benjamins.

- Schwabe, K. & Winkler, S. 2007. *On Information Structure, Meaning and Form: Generalizations across Languages*. Amsterdam: Benjamins.
- Siewierska, A. 1991. *Functional Grammar*. London: Routledge.
- Skopeteas, S. & Fanselow, G. 2009a. Effects of givenness and constraints on free word order. To appear in *Information Structure from different perspectives*, M. Zimmerman & C. Féry (eds). Oxford: Oxford University Press.
- Skopeteas, S. & Fanselow, G. 2009b. Focus in Georgian and the expression of contrast. In *Lingua, Special issue on Contrast*. Ph. Cook & S. Repp (eds).
- Skopeteas, S., Féry, C. & Asatiani, R. 2009. Word order and intonation in Georgian. *Lingua* 119(1): 102-127.
- Skopeteas, S., Fiedler, I., Hellmuth, S., Schwarz, A., Stoel, R., Fanselow, G., Féry, C., & Krifka, M. 2006. *Questionnaire on Information Structure: Reference Manual* (Special Issue in: *Interdisciplinary Studies on Information Structure*, Vol. 4). Potsdam: Universitätsverlag.
- Szendrói, K. 2001. Focus and Syntax Phonology Interface. Ph.D. dissertation, University College London.
- Szendrói, K. 2003. A stress based approach on the syntax of Hungarian focus. *The linguistic review* 20: 37-78.

- Thériault, A., Hellmuth, S. & Skopeteas S. 2008. *The Subject Focus Constraint* and cleft constructions in Québec French. Ms., Universität Potsdam.
- Van Valin, R.D. 1999. A Typology of the interaction of focus structure and syntax. In *Typology and the Theory of Language: From Description to Explanation*, E. Raxilina & J. Testelec (eds). Moscow.
- Vogt, H. 1971. *Grammaire de la langue géorgienne*. Oslo : Universitetsforlaget.
- Wedgwood, D. 2003. *Predication and Information Structure: A Dynamic Account of Hungarian Pre-verbal Syntax*. Ph. D. dissertation, Univ. of Edinburgh.
- Wedgwood, D. 2007. Identifying inferences in focus. In Schwabe and Winkler (eds.), 207-227.
- Wedgwood, D., Pethő, G., & Cann, R. 2006. Hungarian ‘focus’ position and English *it*-clefts: the semantic underspecification of focus readings. Ms., University of Edinburgh.
- Zerbian, S. 2007. The subject/object asymmetry in Northern Sotho. In Schwabe and Winkler (eds.), 323-345.
- Zimmermann, M. 2007. Contrastive focus. In *The Notions of Information Structure* (Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure 6), C. Féry, G. Fanselow, M. Krifka (eds), 147-159. Potsdam: Universitätsverlag.

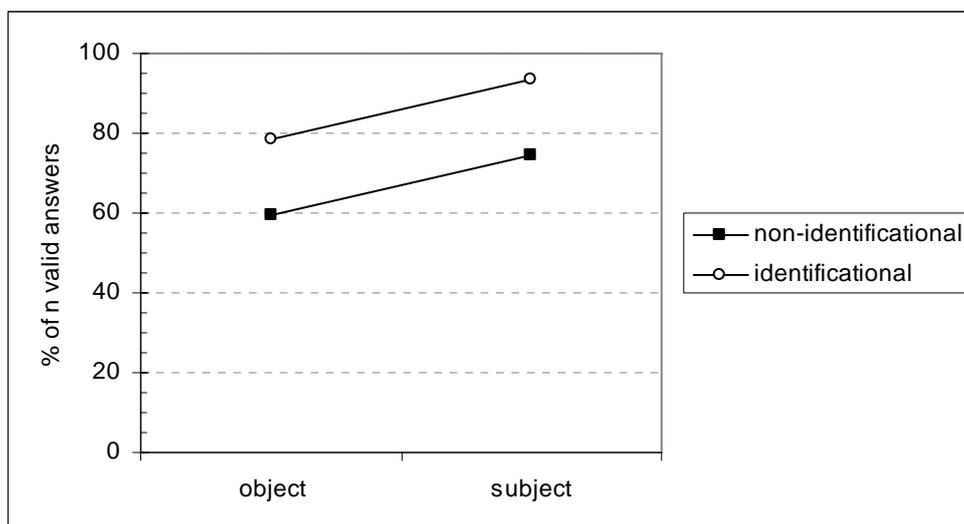
Zimmermann, Malte 2009, Focus realization in West Chadic. To appear in  
*Information Structure from different perspectives*, M. Zimmerman & C.  
Féry (eds). Oxford: Oxford University Press.

### **Glosses**

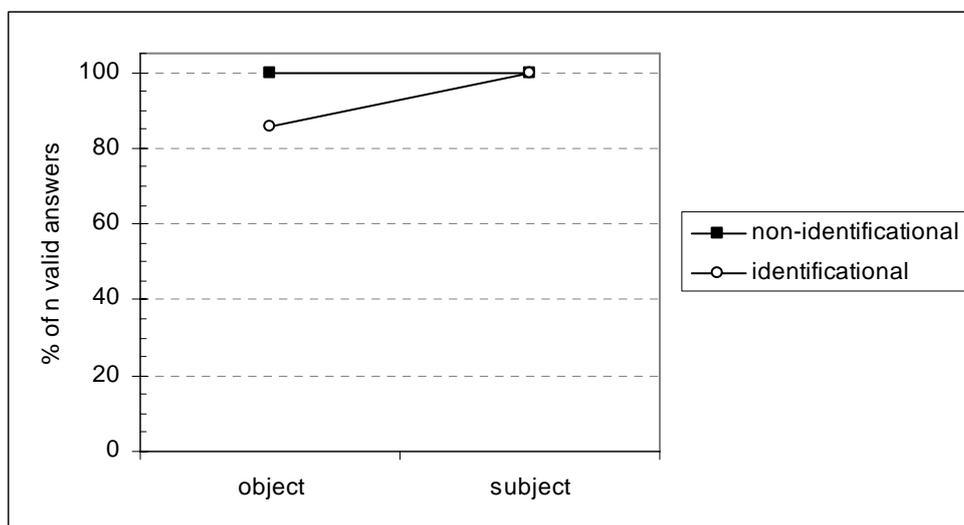
ACC: accusative; DAT: dative; DEF: definite; F: feminine; INDEF:  
indefinite article; IO: indirect object; M: masculine; NOM: nominative; PF:  
perfect; PRF: perfective; PV: preradical vowel; S: subject (person affix);  
SG: singular; SUP: superessive; THM: thematic suffix.

**Figures**

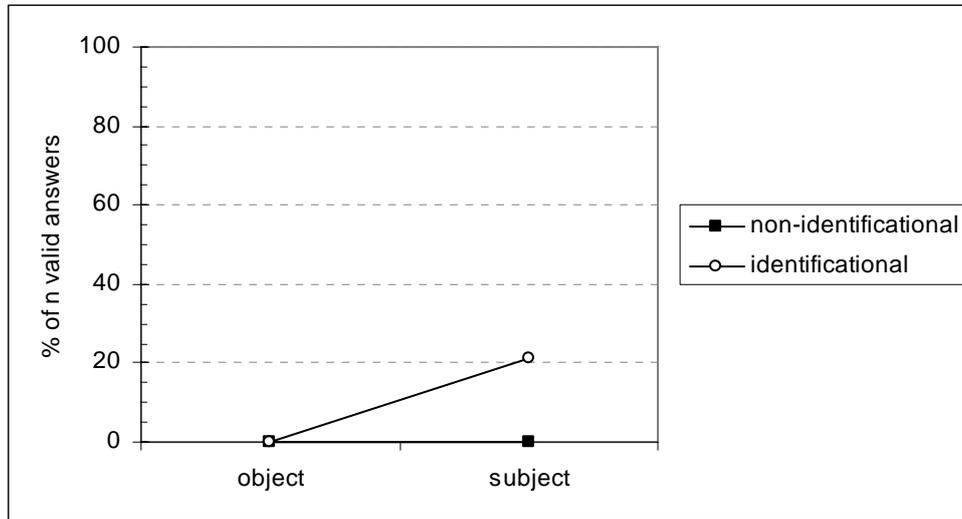
**Figure 1.** Percentage of preverbal focus in Georgian (averages of speakers' means)



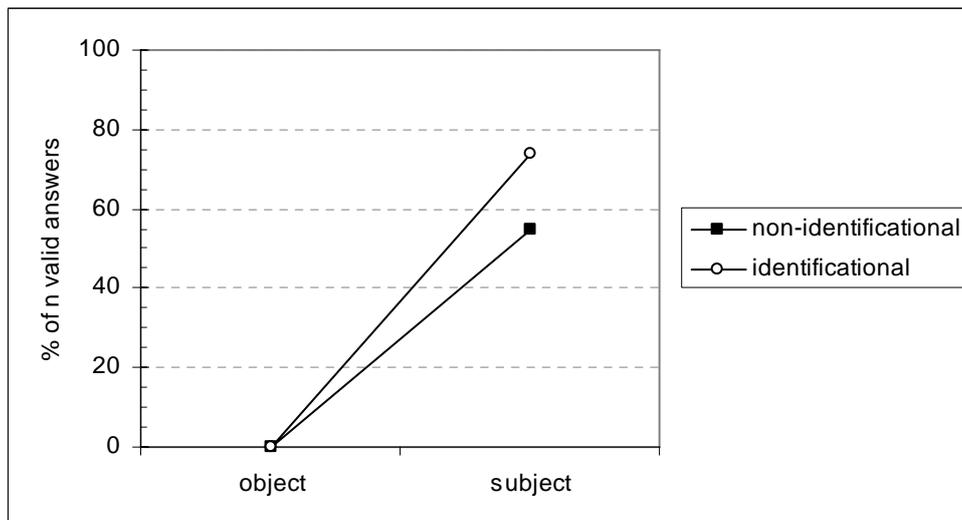
**Figure 2.** Percentage of preverbal focus in Hungarian (averages of speakers' means)



**Figure 3.** Percentage of *it*-clefts in English (averages of speakers' means)



**Figure 4.** Clefted focus constituent in Québec French (averages of speakers' means)



**Tables**

**Table 1.** Georgian data set

	non-identificational				identificational			
	object		subject		object		subject	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
total	40		40		40		40	
non-valid	16		19		14		15	
valid	24	100.0	21	100.0	26	100.0	25	100.0
SVO	12	50.0	11	52.4	7	26.9	20	80.0
SOV	6	25.0	-	-	13	50.0	-	-
OVS	3	12.5	6	28.6	-	-	1	4.0
OSV	-	-	3	14.3	-	-	2	8.0
OV	3	12.5	-	-	6	23.1	-	-
SV	-	-	1	4.8	-	-	2	8.0

**Table 2.** Hungarian data set

	non-identificational				identificational			
	object		subject		object		subject	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
total	8		8		8		8	
non-valid	-		1		1		-	
valid	8		7		7		8	
S pV O	-		-		1	14.3	-	
S Vp O	-		3	42.9	-		1	12.5
SO Vp	2	25.0	-		2	28.6	-	
O Vp S	2	25.0	-		1	14.3	-	
S V O	-		4	57.1	-		7	87.5
S O V	2	25.0	-		3	42.9	-	
O V S	2	25.0	-		-		-	

**Table 3.** English data set

	non- identificational				identificational			
	object		subject		object		subject	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
total	40		40		40		40	
non-valid	12		12		6		12	
valid	28		28		34		28	
simple clause	28	100.0	24	85.7	34	100.0	20	71.4
<i>it</i> -cleft	-		-		-		6	21.4
presentational	-		4	14.3	-		2	7.1

**Table 4.** Québec French data set

	non-identificational				identificational			
	object		subject		object		subject	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
total	80		80		80		80	
non-valid	11		5		10		7	
valid	69	100.0	75	100.0	70	100.0	73	100.0
canonical	68	98.6	34	45.3	68	97.1	19	26.0
cleft	1	1.4	41	54.7	2	2.9	54	74.0
<i>c'est</i>	-		32	78.0	2	100.0	49	90.7
<i>ya</i>	1	100.0	9	22.0	-		5	9.3