Object shift in the Scandinavian languages : syntax, information structure, and intonation

Hosono, M.

Chapter 5. Theoretical Account

5.1. A new hypothesis and generalization on Scandinavian Object Shift

We have long seen that downstep occurs in simple tense forms and Verb Topicalization. By contrast, downstep does not occur in complex tense forms and embedded clauses. OS typically occurs in the former but does not occur in the latter in almost all the Scandinavian varieties investigated, aside from a few exceptions. As shown by the pitch contours presented in chapter 3 and the statistical data presented in chapter 4, it is obvious that there is a strong association between the presence of OS and that of downstep. Thus, the overall property of Scandinavian Object Shift is described as follows: movement of the object pronoun entails downstep. The relation of 'entailment' expresses that whenever OS takes place, downstep occurs (but not vice versa).

There are several reasons for which the entailment relation should be interpreted as a stronger relation. Recall that downstep is typically caused by the L intervening between two Hs; see § 3.1.2. The pitch typically lowers on an object pronoun located between a raised main verb and the negation. This indicates that a weak pronominal object is an inherently low-tone element.[116] The pitch can be high on a shifted object pronoun in some cases – in fact, the pitch peak can occur on it; see the pitch contours of simple tense forms in East Swedish and East and West Norwegian in chapter 3. According to Odden (2007:103), the element that originally has a low tone and plays a role in causing downstep, can appear as a high tone in front of another high-tone element (*H-insertion*) and causes downstep of that following high tone, as illustrated by the second L in the following case: ˩ ˩ ˧ ˩ (L-L-H-L) → ˩ ˧ ꜜ˥ ˩ (L-H-ꜜH (downstepped) -L). Hence, it is not surprising even if an object pronoun that inherently has a low tone appears as a high-tone element in the shifted position and downstep occurs on the following sentential element(s). Note that the fact that a sentential element occurs as a high-tone element does not imply that it is assigned a focal accent. As we have seen, the syllable(s) or word(s) contained in the focal H are not accented, but the pitch level on them becomes higher than that on the accented syllable of a focused word.

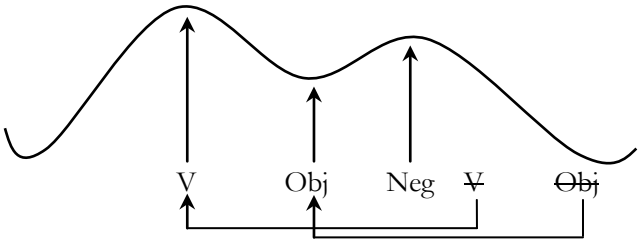
Recall also the data on the non-OS construction of simple tense forms (69-70); see § 3.2.1.1. The pitch does not lower after the accented syllable of the main verb to a considerable extent in this construction. This observation is statistically confirmed. See Appendix III. A-B are the OS construction of simple tense forms with a mono- (A)/disyllabic (B) object pronoun. A'-B' are the non-OS construction of simple tense forms with a mono- (A')/disyllabic (B') object pronoun. The total mean downstep size of A is 2.69; that of A' is

[116] This claim is confirmed by the fact that pitch always lowers on, e.g. short reflexive pronouns.

2.38. The total mean downstep size of B is 2.87; that of B' is 2.39. These data show that regardless of whether a sentence contains a monosyllabic or disyllabic object pronoun, the downstep size is larger in the OS construction than in the non-OS construction. Thus, it is difficult to claim that downstep happens to occur after an object pronoun moves to a surface position. If so, the downstep size would not differ between the OS and non-OS constructions, contrary to fact.

Therefore, the relation of 'entailment', i.e. the presence of OS always coincides with downstep (but not vice versa) should be theoretically interpreted as that of 'causation'. I propose the following new hypothesis on OS:

(165) Scandinavian Object Shift:
The object pronoun moves to cause downstep.



The question is why downstep must be triggered by movement of object pronouns in simple tense forms, i.e. when verb movement takes place. In Swedish, the stressed syllable of an accent 2 word that constitutes the final part of a focal H contour creates the impression that the word itself is focused, as the last high pitch occurs on that stressed syllable (Gussenhoven 2004:213). The negation *inte* is an accent 2 word. If the first syllable *in-* of the negation were to carry a focal H contour as its final part, it might sound as if the negation itself were focused. However, the focus of a sentence is carried by a raised main verb and the focal accent of a sentence is located on it in the relevant context (of polarity-focus), as mentioned in § 3.2.1.1. Hence, an object pronoun moves, causes downstep, and prevents a focal H contour from arising after an accented main verb. In other words, movement of object pronouns serves to eliminate a focal effect on the negation on one hand and to maintain the focus of a sentence on the main verb on the other.

We also saw the case in which a focal H contour actually occurs in the OS construction of simple tense forms. Here too, movement of an object pronoun makes the F0 on the negation lower than the F0 on an accented main verb, i.e., prevents a focal effect from arising on the negation, with the object pronoun itself being a high tone element and actually the pitch peak.

The argument above also applies to Verb Topicalization, where the pitch does not rise again after it falls on a sentence-initial past participle. Since a contrastively focused past participle is the sole possible locus of the focus of a sentence, an additional focus of a sentence that could be realized by a focal H contour must not occur: a sentence can have one and only one focus (Lambrecht 1994); see § 3.1.1. Thus, an object pronoun moves and triggers downstep to prevent a pitch rise for a possible focal H contour that could produce a focal effect on the negation from arising.[117]

In Norwegian, when the negation follows a main verb, they compose a LH pitch contour together. This indicates that after the accented syllable of a main verb, the first syllable of the negation becomes an accented L, i.e. L*, and the H occurs on its second syllable. The H that would be expected to occur on the second syllable of a main verb disappears. This situation is allowed when a focused word follows the unit consisting of a main verb and the negation: the focal H occurs on the H on the focused word that follows that unit. In the case in which a sentence has only a main verb, the negation and an object pronoun, however, the focus and focal accent of a sentence is carried by a main verb. The focal H must occur in the pitch contour of a main verb. If a main verb and the negation composed a unit, the first syllable of the negation would be an accented L*, and the focal H would occur on its second syllable. The negation itself then might sound focused. Hence, an object pronoun moves, causes downstep and eliminates a focal effect on the following negation.

In (East) Danish, when a monosyllabic object pronoun follows a main verb, stød occurs on the second, unstressed syllable of the latter. The question is what relationship there is between the occurrence of stød on (the second syllable of) the main verb and the presence of movement of the (monosyllabic) object pronoun. The pitch is relatively high on the second syllable of the main verb. Recall that when the negation directly follows the main verb, it composes an independent accentual unit and is assigned a prominence. It would be expected that the pitch continues to rise up to the negation, and the peak occurs on it. This might sound as if the negation were focused. However, the focus and focal accent of a sentence is carried by the main verb. Therefore, an object pronoun moves and causes downstep to eliminate a focal effect on the negation. When a monosyllabic object pronoun moves, the preceding high-tone syllable (of the main verb) becomes the pitch peak and a stød occurs on it: movement of an object pronoun triggers the occurrence of stød. This movement further triggers the pitch lowering on the

[117] Note that an object pronoun moves to eliminate a possible pitch rise, not a possible focal accent, on the negation. When the negation is contained in the focal H, the focal effect on it could be produced by the highest tone on its first syllable, rather than by the accent on it. Recall the statement in § 3.2.1: the two-peaked varieties, e.g. East Swedish, maintain a word accent on non-focused words; thus, after the point of the focal H, downstep of the following non-focal accents occurs to produce a difference in pitch level between focal and non-focal accents.

following sentential element(s). The pitch lowers partly due to the inherent low status of an object pronoun and partly due to the general property of stød that the pitch drastically lowers after it. When a disyllabic object pronoun moves, it directly causes downstep due to its inherently low-tone status.

In Icelandic, a main verb, a shifted object pronoun and the negation compose a phonological word in the cliticization process in the OS construction of simple tense forms. The pitch peak occurs on the main verb that carries the focus and focal accent of a sentence. Thanks to movement of an object pronoun, the resulting word order, V+Obj_{pro}+Neg, produces a phonological word in which only the main verb is accented and the negation does not attract a secondary stress. In other words, an object pronoun contributes to lowering the pitch level on the negation significantly in the typical OS construction by moving and triggering the birth of a phonological word composed of the main verb, the shifted object pronoun itself and the negation in that order. Thus, movement of an object pronoun causes downstep on the following negation.

In Faroese, the pitch peak normally occurs on an accented main verb, and the pitch drastically falls after it. The pitch continues to be low until sentence-final position. When the negation intervenes between a main verb and an object pronoun, the pitch rises on the sentence-final object pronoun, as illustrated in the case of embedded clauses. This indicates that the pitch does not lower on the negation preceding the object pronoun. See (156). The pitch falls after the (embedded) main verb, but the pitch level on the negation is relatively high compared with the second syllable *-na* of the object pronoun in sentence-final position. Thus, the object pronoun moves to trigger downstep and lower the pitch level on the negation.[118]

In complex tense forms and embedded clauses, in which both main verb movement and OS typically do not occur, on the other hand, the pitch peak occurs either on the negation, on a main verb, or on an object pronoun (in situ). In the relevant context (of polarity-focus), the focus and focal accent of a sentence is carried by the in-situ past participle main verb located after the Aux and by the (normally in-situ) embedded main verb located after the embedded subject. The final pitch peak occurs on those main verbs. Therefore, an object pronoun must not move and cause downstep before them.[119] This argument applies to all the Scandinavian varieties investigated here.[120]

---

[118] Recall that when an object pronoun is isolated from an embedded main verb, a sentence sounds as if an object pronoun were somewhat prominent. Thus, it could be argued that an object pronoun moves to avoid a high tone on itself. However, the pitch can become high on an in-situ object pronoun. See, e.g. the complex tense form with a monosyllabic object pronoun (153b). Thus, there is no reason for which an object pronoun has to move to avoid a high tone on itself.

[119] As we have seen, the final pitch peak can occur on an in-situ object pronoun.

[120] This argument also accounts for the fact that an object pronoun cannot directly follow the negation either:

The account of Holmberg's Generalization is thus provided as follows. When main verb movement takes place, an object pronoun moves and causes downstep to eliminate a focal effect on the sentential element(s) after the main verb. In the environments in which downstep must not occur, i.e. in the constructions where the final pitch peak occurs on the (in-situ) main verb, OS does not occur either.

Note that the argument here straightforwardly accounts for the optional aspect of OS. That is, OS being optional, an object pronoun does not need to move for its own reason at all. Why does it still move? The reason is that it moves to trigger downstep and eliminate a focal effect on the sentential element(s) following the main verb.

With the argument made so far, the cases below are accounted for. As introduced in 2.1, an object pronoun cannot move across an indirect full NP object (166a). It (normally) cannot move across a subject in *yes-no* questions either (166b) (Holmberg 1986, 1999). In addition, it generally cannot move across a main verb (166c).[121] In some *pro*-VP forms (166d) and some copula sentences (166e), an object pronoun cannot move.

(166) a. Jag gav inte Elsa *den*. *Jag gav *den* inte Elsa. [Swe.]
I gave not Elsa it I gave it not Elsa
'I didn't give it to Elsa.'

b. Köpte Johan *den* inte? *Köpte *den* Johan inte?
bought Johan it not bought it Johan not
'Didn't Johan buy it?'

c. Jag köpte den inte. *Jag *den* köpte inte.
I bought it not I it bought not
'I didn't buy it.'

d. ('You in fact slapped him, didn't you do that?')
– Nei, jeg gjorde ikke det. [Nor.]
no, I did not it
'No, I didn't do so.'

e. (The fastest player in the team is without doubt Morten and)
den hojeste er (*ham) faktisk også (OKham). [Dan.]
the tallest is him actually also him
'the tallest one/player is actually also him.'

(i) *Jag har inte den malat. [Swe.]
I have not it painted
Since the final pitch peak occurs on the past participle, an object pronoun must not move across it and cause downstep before it.

121 Thanks to Anders Holmberg (p.c.) for pointing out (166c) to me.

In appropriate contexts, the indirect full NP *Elsa* (166a), the subject *Johan* (166b), and the main verb *köpte* (166c), are the most appropriate candidates for the carrier of the focus in each of the sentences. The final pitch peak occurs on those focused elements. Since downstep must not occur before them, an object pronoun must not move and cause downstep; thus it cannot cross them.[122] Recall that the answer sentence in (166d) is a denial of the proposition presented in the preceding sentence. That is, the negation carries the focus of the answer sentence. Thus, the pitch must rise up to the negation; an object pronoun must not move to trigger downstep before the negation. As mentioned in § 2.2, the post-copular domain is focused in (166e). The pitch must not lower before the adverbs in the focused, post-copular domain due to movement of an object pronoun.[123]

In chapter 3, a new generalization on OS was presented:

(167) Generalization on Scandinavian Object Shift (second approximation):
The more delayed the pitch gesture is, the more likely is Object Shift to be absent in a relevant Scandinavian variety.

The formulation above is based on the fact that Övdalian and Dalecarlian in general, in which OS tends to be absent, have a delayed pitch gesture compared with the other Swedish varieties. Depending on whether the pitch gesture occurs in an early or late timing, on one hand, and whether the dialect at issue is one- or two-peaked, on the other, the Swedish varieties (and Mainland Scandinavian in general) are classified into the following three types: i) early timing (HL) and two-peaked – e.g. East Swedish; ii) late timing (LH) and one-peaked – e.g. South Swedish and Dalecarlian; and iii) late timing (LH) and two-peaked – Övdalian. OS is more or less obligatory in i), optional in ii) and totally absent in iii).

---

122 An object pronoun can in fact move across a main verb, when it is contrasted:
(i) Den köpte jag inte. [Swe.]
it bought I not
'I didn't buy IT (, but THAT).'
The construction above can be used as either focalization or topicalization of an object pronoun, which has not been called OS. But in this kind of construction, the pitch peak occurs on the contrasted object pronoun in sentence-initial position. The pitch lowers after it and does not rise again. The fact that an object pronoun can be contrasted does not contradict the fact that it is an inherently low-tone element: as long as it is an argument of a verb, it can always be contrasted. Thanks to Johan Rooryck (p.c.) for pointing out this to me.

123 As we have seen, an object pronoun cannot move across a verb participle in Swedish and Övdalian but moves in the other Scandinavian varieties; see § 2.1. Anders Holmberg (p.c.) suggests the possibility that such a difference in word order is a matter of a parameter concerning VP-internal linearization. In my paper (in prep.), I show that the difference in word order among the Scandinavian languages is closely related to the intonational properties of verb particle constructions.

The only gap is the type of the Scandinavian varieties that have an early pitch gesture (HL) and are one-peaked. The Insular Scandinavian varieties, Icelandic and Faroese, do not have word accents. Word stress almost always occurs on the first syllable. Recall that though Icelandic can have various kinds of pitch accent, the LH contour cannot appear in neutral declarative sentences unless it is followed by another H*. This indicates that Icelandic has H*L(*) as the basic pitch gesture. The Icelandic pitch accent system is very similar to, e.g. that of English, and phrasal accent is located on the rightmost constituent in the unmarked case. In other words, the pitch peak occurs only once on the most prominent (rightmost) word of a sentence in Icelandic. All of these intonational properties are shared by Faroese. That is, though Icelandic and Faroese do not have word accents, they can be regarded as having the properties of the pitch gesture with an early timing (HL) and one-peaked.

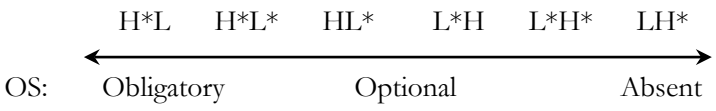
With the final gap filled by Insular Scandinavian, the Scandinavian varieties are classified as follows:

(168) Classification of the Scandinavian varieties:

	HL (early timing)	*LH* (late timing)
one-peak	Ice., Far.	Swe. (South, North, Fin.-Swe., Dal.) Nor. (West), Dan. (East, South)
two-peak	Swe. (East, West) Nor. (East)	Swe. (Övd.)

The relationship between the early/late pitch gesture and the obligatoriness/optionality/absence of OS is illustrated in (168). Below, OS is more obligatory in a Scandinavian variety that has the pitch gesture located more to the left. OS is more optional and even absent in one that has the pitch gesture located more to the right.

(169) Relationship between early/late pitch gesture and obligatoriness/ optionality/absence of Object Shift:



Whether OS is obligatory, optional or absent depends on whether a Scandinavian variety at issue has an early or late pitch gesture. Hence, the generalization (167) is reformulated as follows:

(170) Scandinavian Object Shift (final):
The earlier the pitch gesture occurs, the more likely is Object Shift to occur; the more delayed the pitch gesture is, the more likely is Object Shift to be absent.

The generalization above accounts for not simply the presence or absence of OS. It also accounts for the extent to which OS occurs, i.e. its obligatoriness, optionality and absence. The extent to which OS occurs depends on the timing of the pitch gesture in each of the Scandinavian languages. OS is then a gradient phenomenon rather than a binary/dichotomous property in the Scandinavian languages.

Let us see the extent to which OS is obligatory, optional and absent in each of the Scandinavian languages, by considering the (un)acceptability of the non-OS construction of simple tense forms A'-B' and comparing the total mean grades of A'-B'. See Appendix II. The total mean grade in East and West Swedish is 2.9, and the one in all the other Swedish varieties is 3.4. This indicates that this construction is less acceptable in East and West Swedish, in which the pitch gesture occurs early, than in all the other Swedish varieties in which the pitch gesture is delayed.

The mean grade of the non-OS construction of simple tense forms in Övdalian is 3.9. Thus, this construction is acceptable for speakers of Övdalian, which has a delayed pitch gesture but are two-peaked.

The total mean grade of the non-OS construction of simple tense forms in East Norwegian is 3.1, and the one in West Norwegian is 2.8. This indicates that this construction is more acceptable in the former than in the latter. The intonational properties in East Norwegian are similar to those in the two-peaked varieties, and the intonational properties in West Norwegian to those in the one-peaked varieties. However, the former has LH as the basic tone, and the latter has HL as the basic one. It is plausible that the difference in the basic tonal pattern is related to the different grade in the acceptability of this construction.

The total mean grade of the non-OS construction of simple tense forms in East Danish is 2.5, and the one in South Danish is 2.6. Both grades are lower compared with the other Mainland Scandinavian varieties, though none of the Danish speakers marks this construction as ungrammatical. The result that this construction is slightly less acceptable in the former than in the latter is compatible with the observation in the literature that OS can be more optional in South Danish than in East Danish (cf. § 3.2.3.2).[124]

The total mean grade of the non-OS construction of simple tense forms in Icelandic is 1.2, and the one in Faroese is 2.6. The grade in Icelandic is significantly lower compared with the other Scandinavian varieties. Almost all

[124] I turn to the question why OS is more or less obligatory in East Danish compared with South Danish below soon.

Icelandic informants mark this construction as ungrammatical. The grade in Faroese is the same as that in the Danish varieties, and low compared with the other Scandinavian varieties.

With the hypothesis (165) as well as the generalization on OS (170), all the aspects of OS, i.e. the relationship between the early/late pitch gesture, the one-/two-peaked property, and the obligatoriness/optionality/absence of OS, are accounted for as follows (cf. Bruce 2005, 2007; see § 3.2.1):

i) The Scandinavian varieties in which the pitch gesture occurs early are typically the two-peaked dialects. In these varieties, the two-level difference in prominence between a focused word and a non-focused word as well as the distinction in word accents are maintained. Downstep should occur to differentiate the two prominence levels. When a sentential adverb follows an accented main verb, however, the point at which downstep could occur is after the primary stressed syllable of a sentential adverb, e.g. after the first syllable *in-* of the negation *inte*, on which the final peak of the focal H would occur. Hence, an object pronoun needs to move and cause downstep to eliminate a focal effect on the negation. OS is more or less obligatory in the relevant contexts in these dialects.

ii) The Scandinavian varieties in which the pitch gesture occurs relatively late are typically the one-peaked dialects. Since the pitch gesture of an accented word always overlaps that of a focal H tone, the prominence level is simply either accented or not. Deaccentuation of non-focused words then occurs instead of downstep. Deaccentuation indicates that the difference in word accents is lost: neutralization of word accents occurs. Neutralization in the context of OS means that the negation, an accent 2 word, loses its H after an accented main verb. Since the pitch has already lowered, an object pronoun does not need to move to cause downstep. OS is more likely to be optional or even absent in these varieties.[125] Recall that deaccentuation occurs only when it is especially required. In the cases in which it is difficult to induce it, e.g. in the simple tense form with a monosyllabic weak pronominal object, OS is more likely to occur to trigger downgrading.[126]

[125] Recall that South and Finland Swedish and Dalecarlian speakers tend to prefer an in-situ object pronoun, especially for disyllabic object pronouns. See Appendix II.

[126] In Danish, the H of the words surrounding a focused word is downgraded. But that occurs only when the focused word has a low vowel. Thus, in (South Danish) *jeg så ikke det* (I saw not it 'I didn't see it'), an accented main verb has a low vowel [ɔ]. OS can be optional here. When an accented main verb does not have a low vowel, OS is more likely to occur.

iii) Övdalian belongs to the two-peaked dialect group, but the pitch gesture occurs late. Unlike in the typical two-peaked varieties in which the focal H starts immediately after an accented syllable of a focused word, the starting point of the focal H is delayed. It occurs at earliest on the next accentable syllable, since the pitch always lowers on an accented syllable of a focused word in the two-peaked varieties. The (final part) LH of LH*LH of a preceding word always overlaps the (first part) LH* of the following word. Since the (second) H of a preceding word always overlaps the (first) H* on the following word, downstep does not occur in Övdalian. An object pronoun cannot, and must not, move to cause downstep. OS is totally absent in Övdalian.

iv) The Insular Scandinavian languages, Icelandic and Faroese, have been classified as belonging to the varieties that are one-peaked but have an early pitch gesture. Downstep could basically occur after an accented main verb. But the pitch does not lower on the negation when it directly follows a main verb, as we saw in Faroese embedded clauses. Thus, an object pronoun needs to move to trigger downstep. In the same way as in the one-peaked varieties, the pitch level is distinguished simply by being accented or not. But these varieties basically have the trochaic accent system. Since deaccentuation of non-focal words does not always occur, the first syllable of each sentential element could have the same prominence. Thus, there is more obligation for an object pronoun to move to cause downgrading in these varieties than in the typical two-peaked varieties with an early pitch gesture.

In § 2.1, we saw the issues on OS classified into the following three items: i) the presence (and absence) of movement of various kinds of pronominal forms; ii) parametric differences among the Scandinavian languages; and iii) particular syntactic properties of OS. We also saw that none of the semantico-syntactic, syntactic and phonological approaches succeeds in providing an account for all of them. Here, item i) is accounted for as follows: an (object) pronoun moves to trigger downstep; OS is absent when downstep must not occur. Item ii), i.e. the obligatoriness, optionality and absence of OS, is accounted for in terms of the intonational properties of each of the Scandinavian varieties we have seen so far.[127] Item iii) is straightforwardly accounted for in terms of the fact that OS occurs for purely intonational/phonological reasons.

I mention several issues in the rest of this section. It has long been argued that OS is more or less obligatory in East Danish. Both East and South

[127] Regarding the parametric difference in verb particle constructions between the Scandinavian languages, see footnote 123.

Danish have LH as the basic tone. But the target H typically occurs on the syllable next to an accented syllable in East Danish (cf. § 3.2.3.1), which is not a typical feature of South Danish (cf. § 3.2.3.2). As stated previously, when the negation directly follows the main verb, it composes an independent accentual unit and is assigned a prominence. Thus in East Danish, if the second syllable of an accented verb on which the target H occurs were immediately followed by the negation, the pitch would continue to rise up to the negation and a focal effect could be produced on it. In South Danish, on the other hand, the pitch can lower on the accented syllable of a main verb, and does not directly rise up to the negation. This will reduce a possible focal effect on the negation. Hence, movement of an object pronoun to cause downstep is more necessary in East Danish than in South Danish.

A full NP interpreted as specific/given can move in Icelandic and is subject to Holmberg's Generalization; see § 2.1:

(171) Jón hefur (*þessa bók) aldrei lesið (^{OK}þessa bók). [Ice.]
Jón has this book never read this book
'Jón has never read this book.'

Above, when the definite NP *þessa bók* moves, the sentence is ungrammatical. Recall that in a definite NP, the leftmost constituent is accented. Thus, the leftmost word *þessa* should be accented in the shifted (and in-situ) position. The pitch level on it should get high. The pitch level on the Aux is also high, as we have seen so far. A possible account for this fact is that there is an intonational clash between the high pitch on the Aux and the high pitch on the leftmost constituent of the full NP following and adjacent to it after movement of the NP.[128]

Finally, it is predicted that OS would not occur when the pitch peak occurs on the sentential element following an object pronoun. There are in fact cases in which the pitch peak can occur on the sentential element following a shifted object pronoun: e.g. contrastive argument-focus of a direct object in ditransitive constructions. In the Scandinavian languages, the direct object normally follows the indirect object. Below, the direct object *en bok* 'a book' is contrastively focused; the pitch peak is likely to occur on it. But the indirect object pronoun *henne* moves across the negation *inte*.

(172) Jag gav henne inte en BOK (, men en KAKA). [Swe.]
I gave her not a book but a cake
'I didn't give her a BOOK (, but a CAKE)'.

[128] A bare nominal can move in simple tense forms when it receives a specific interpretation. But in complex tense forms, an intonational clash would occur, since Icelandic has a trochaic accent system. The high pitch on the Aux and the high pitch on the first syllable of the shifted nominal would be adjacent to each other.

According to Holmberg and Platzack (1995), a dative object is exceptional in that not only a pronominal form but also a full NP can move across a sentential adverb:

(173) a. De ga Marit ikke blomstene. [Nor.]
they gave Marit not the-flowers
'They didn't give Marit the flowers.'

b. Vi ger barnen alltid vad de vill ha. [Swe.]
we give the-children always what they want have
'We always give the children what they want to have.'
(Holmberg and Platzack 1995:172, (6.60b-c))

Thus, movement of the indirect object pronoun in (172) is not unexpected. It can occur due to the general property of dative objects described in (173).[129,130]

5.2. Interaction between syntax, information structure and intonation

In this section, I propose a new system that accounts for the facts regarding OS as well as the interaction between syntax, information structure and intonation in general, which is not based on any existing theoretical framework. *Information structure*, the term originated from Halliday (1967), has long been studied in different disciplines and frameworks (Mathesius 1929; Firbas 1974; Chomsky 1970; Jackendoff 1972; Gundel 1974; Chafe 1976; Kuno 1976; Li and Thompson 1976; Dik 1978; Givón 1979; Selkirk 1984; Vallduví 1990; Vallduví and Engdahl 1996; Lambrecht 1994; Rizzi 1997; Zubizarreta 1998; Bresnan 2001; Steube 2004; Hengeveld and Mackenzie 2006; Schwabe and Winkler 2007, among others). In this thesis, I define information structure as the discourse concepts that mediate between grammatical components such as syntax and phonology to express the information flow of a sentence in a language at issue. The basic concept is focus, the center of a given discourse, which plays a central role in the system proposed below. For the theory of

---

129 The answer to the question why dative objects behave in an exceptional way is beyond the topic of this thesis.

130 Anders Holmberg (p.c.) points out that the negation in simple tense forms can be contrastively focused and the pitch peak can occur on it:

(i) Jag målade den INTE. [Swe.]
I painted it not
'I did NOT paint it.'

In fact, the prominence occurs on the negation not only in the context of contrastive focus. The pitch peak occurs on it also in the context of polarity-focus for some speakers. The prominence on it can occur depending on speakers' preference. The speaker who tends to raise the pitch on the negation in simple tense forms also tends to do so in other constructions. I leave the fact that the negation is prominent (in the contexts of polarity-focus/contrastive focus) but OS can occur for future research.

sentence accentuation, see Chomsky and Halle (1968), Schmerling (1976), Gussenhoven (1984), Selkirk (1984, 1995), Rochemont (1986), Cinque (1993), Zubizarreta (1998), Kahnemuyipour (2009), among others.

We have seen that the highest pitch peak occurs on the accented syllable which is contained in the focal H that occurs immediately after a focused word in the Scandinavian two-peaked varieties. It overlaps the H of the pitch gesture of a focused word in the one-peaked varieties. Thus, the locus of the highest pitch peak always indicates that the focal point is also there (or quite near it). The locus of an accent does not necessarily show the focal point, however: both a main verb and the negation are accented in the OS construction of simple tense forms, but only the former, not the latter, carries the focus of a sentence. That is, the locus of an accent is not primary unlike that of the pitch peak.

The basic idea is that in theorizing the interaction between syntax, information structure and intonation, only the focal point and the highest pitch peak point need to be taken into account. Since the highest pitch peak point always points to the focal point on it or quite near it, the former must more or less correspond to the latter. This is formulated as the following principle:[131]

(174) The highest pitch peak point mostly coincides with the focal point.

As stated so far, a sentence can have one and only one focus, though it can have more than one non-focused/topic element (Lambrecht 1994, Rizzi 1997). Therefore, the location of the focus of a sentence should play the central role in theorizing the interaction between syntax, information structure and intonation.

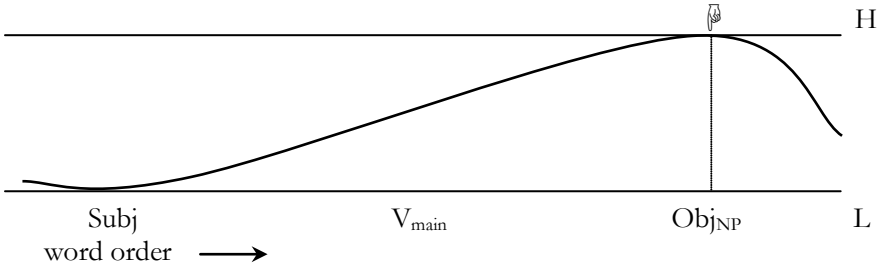
In the proposed system, the focal point of a sentence is indicated by the focal pointer *fp*, ' ☞ '. The fp is the indicator of the change in the information flow of a sentence. When it moves, the focal point (and the pitch peak) moves too. Below, the syntactic word order, here SVO, goes on from the left to the right. In transitive constructions, the focus is carried by a (full NP) object in the unmarked case (Gundel 1988). The fp is normally located on it. The pitch rises towards the object, and the pitch peak occurs on its rightmost constituent. The pitch peak coincides with the focal point indicated by the fp. After the pitch peak, downstep occurs in the position(s) following the focal point.[132]

---

[131] This principle is compatible with the widely claimed view in the literature given above: the focused constituent must contain the word most prominent in a sentence.

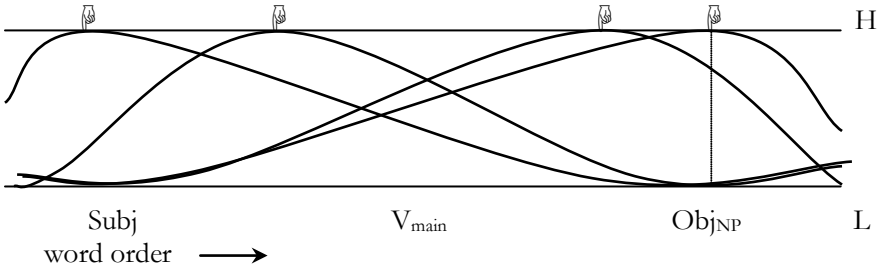
[132] Abbreviations: H – high; L – low; Subj – a subject; S.Adv – a sentential adverb; Aux – an auxiliary verb; V_{main} – a main verb in a main clause; V_{part} – a past participle; V_{emb} – a main verb in an embedded clause; Obj_{NP} – a full NP object; Obj_{pro} – an object pronoun.

(175) The unmarked case of transitive constructions in SVO languages:



When the fp occurs on the unmarked position, the unmarked intonation pattern does not change; the syntactic word order is not affected either. As the fp moves from the right to the left, however, the focus of a sentence occurs on a marked position. The pitch peak moves too and occurs on the constituent in that marked position. Since downstep starts from the marked position, the unmarked intonation pattern changes. The syntactic word order is also likely to be affected, often by movement of a sentential element. Thus, as illustrated below, the farther to the left the fp moves from the unmarked object position, the more the unmarked intonation pattern is likely to change, and the more the unmarked SVO order is likely to be affected.

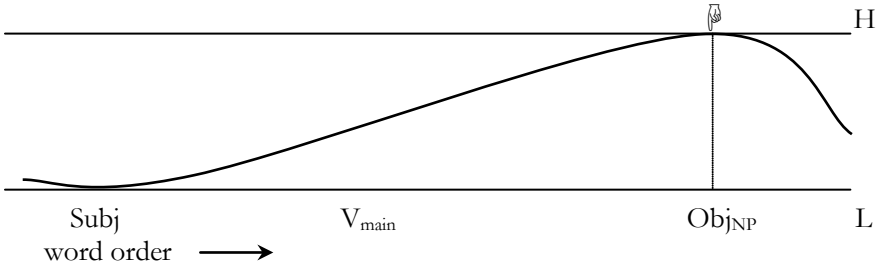
(176) The interaction between the changes of the focal point, the pitch peak point and the syntactic word order in SVO languages:



The facts on OS we have seen so far are accounted for as follows. The Scandinavian languages are SVO languages. Since the focus is carried by an object, the fp is located on the sentence-final full NP object in the unmarked case (regardless of whether a main verb moves and/or whether a sentential adverb is present) (177a).[133] The pitch rises towards that focal point. The focal H occurs on the object that carries the focus, and the pitch peak occurs on it. The unmarked intonation pattern does not change. The unmarked syntactic word order is not affected either. Some examples are given in (177b).

[133] This case includes that of a focused object pronoun in situ.

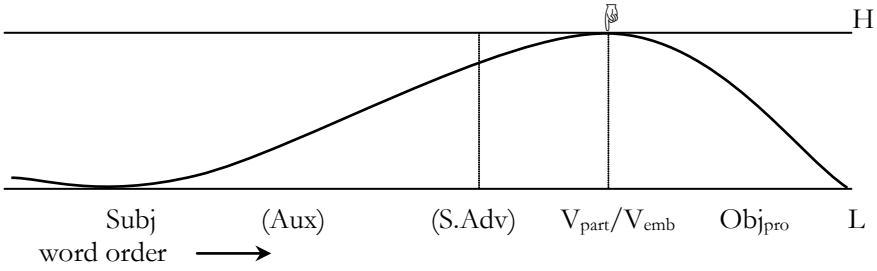
(177) a. Unmarked cases:



b. Jag kysste *Anna.*[134] 'I kissed Anna.' [Swe.]
Jag har sett *filmen.* 'I have seen the movie.'
Jag sa att jag kysste *Anna.* 'I said that I kissed Anna.'

In complex tense forms and embedded clauses that have a non-focused object pronoun, the focus of a sentence is carried by a past participle/embedded main verb located (mostly) in situ. The fp moves from the object position and occurs on it (178a). The final pitch peak occurs on it too (regardless of the presence or absence of a sentential adverb). Since downstep must not occur before it, OS does not occur either. Thus, the unmarked intonation pattern does not change; the unmarked syntactic word order is not affected either. Some examples are given in (178b).[135]

(178) a. Complex tense forms and embedded clauses with a weak pronominal object:



[134] The locus of the (information/contrastive) focus of a sentence is indicated by italics.

[135] The case in which Holmberg's Generalization is violated (i) is derived from the violation of principle (174). If the object pronoun moved to cause downstep, the pitch peak point could not occur on the past participle main verb and not coincide with the focal point on it. Thus, it must not move.

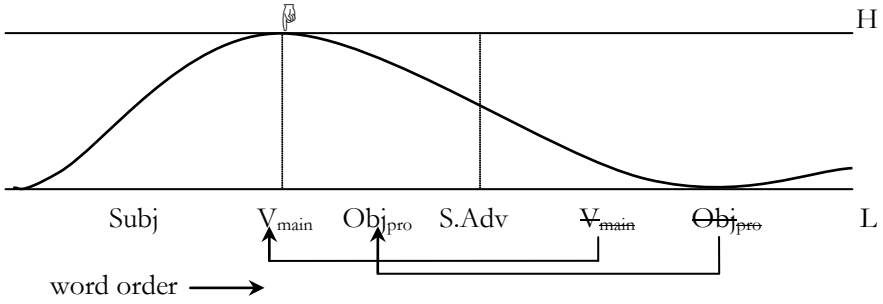
(i) *Jag har honom inte målat. [Swe.]
I have him not portrayed

Thanks to Anders Holmberg (p.c.) for pointing out this to me.

b. Jag har *sett* den./Jag har inte *sett* den. [Swe.]
'I have seen it/I have not seen it.'
Jag sa att jag *kysste* henne./Jag sa att jag inte *kysste* henne.
'I said that I kissed her/I said that I not kissed her.'

In simple tense forms that have a non-focused object pronoun and a sentential adverb, the focus is carried by a main verb. The fp moves and occurs on the main verb in the second position (179a). The pitch peak occurs on it too. After that pitch peak, the pitch must lower. Hence, the object pronoun moves and causes downstep to prevent a focal effect from arising on the sentential adverb located after the main verb. Since the pitch peak occurs on the raised main verb and downstep starts immediately after it, the unmarked intonation pattern changes. The unmarked word order is also affected as illustrated by the presence of OS. (179b) is some illustrations. Note that if OS did not occur, the pitch peak could occur on the negation, as discussed in § 5.1. This yields the situation against principle (174): since the focal point would occur on the main verb but the pitch peak on the negation, the focal point would not coincide with the pitch peak point.

(179) a. Simple tense forms with a weak pronominal object:

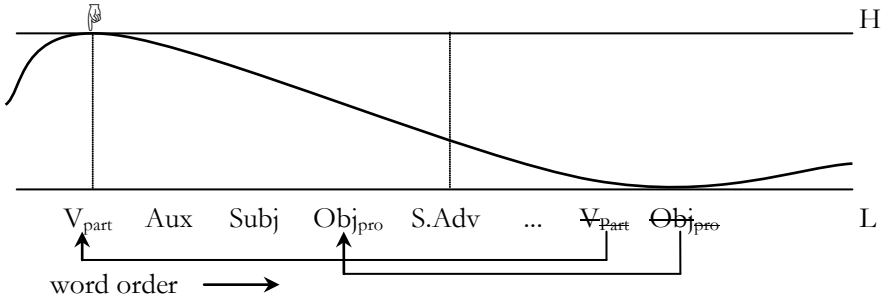


b. Jag *kysste* henne inte. 'I kissed her not.' [Swe.]
Jag *köpte* den inte. 'I bought it not.'

In Verb Topicalization, the past participle main verb is contrastively focused. The fp moves farther to the left and occurs on the past participle in sentence-initial position (180a). The pitch peak occurs on it too. After that pitch peak, downstep must occur. An object pronoun moves and causes downstep to prevent a focal effect from arising on the sentential adverb that has the next accentable syllable. Since downstep starts immediately after the pitch peak on the sentence-initial past participle, the unmarked intonation pattern changes further than in the cases we have seen above. The unmarked syntactic order is also more affected than in the cases so far, as illustrated by the presence of past participle fronting resulting in the VSO order and that of OS.

(180b) gives an illustration.

(180) a. Verb Topicalization:



b. *KYSST* har jag honom inte. 'I haven't KISSED him.'

The cross-linguistic prediction from the model above is as follows: the farther the fp moves from an unmarked position, the more an unmarked intonation pattern is likely to change, and the more an unmarked syntactic word order is likely to be affected.

In SVO languages, the focus of a sentence and the final pitch peak occur near the end of the sentence, e.g. on the (full NP) object, in the unmarked case; the fp occurs on the element located there. See (175) for the illustration of the unmarked case. It is predicted that the farther the fp moves to the left, the unmarked intonation pattern is more likely to change; the unmarked syntactic word order is more likely to be affected too. This is confirmed by the case of object argument-focus. The focus of a sentence is carried by the *wh*-phrase *what* in (181a) and by the full NP *that article* in (181b). The fp and the pitch peak move from the original position to sentence-initial position as illustrated in (181c). Since downstep occurs immediately after those focused elements in sentence-initial position, the unmarked intonation pattern changes. The syntactic word order is affected too, as illustrated by the presence of *wh*-movement (181a) and focus fronting (181b).[136]

---

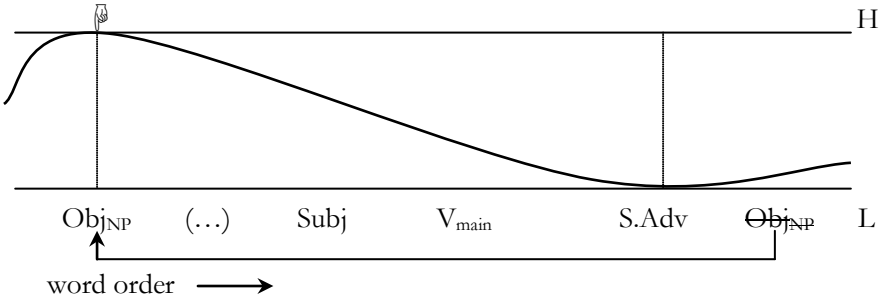
136 An interesting case is Hungarian, an SVO language. According to Szendrői (2003), a focal accent occurs in a sentence-medial position in Hungarian. It occurs on a main verb in broad-focus contexts, as illustrated in (i), where the accent on the unit composed of the particle and the main verb *kinézett* 'chose' is more prominent (indicated by large capitals) than that on the object *kalapot* 'hat'. In other cases, the focal accent occurs strictly in the immediately preverbal position. Thus, a focused sentential element moves and immediately precedes the main verb, as illustrated by the object *kalapot* in (ii).

(i) Mari KINÉZETT magának egy KALAPOT. [Hun.]
Mari PRT-spotted herself-DAT a hat-ACC
'(What did Mary do?) Mari chose a hat for herself.'

(181) a. *What* did you do yesterday*?*

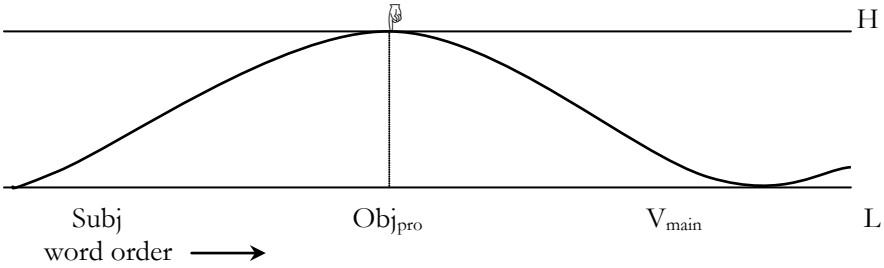
b. *THAT ARTICLE*, I didn't read today.

c.



In SOV languages as well, the focus of a sentence is carried by a (full NP) object in transitive constructions in the unmarked case (Gundel 1988). The fp and the pitch peak occur on the position immediately preceding a verb. The unmarked case is illustrated below:

(182) The unmarked case of transitive constructions in SOV languages:



It is predicted that the farther the fp moves either to the left or to the right, the more the unmarked intonation pattern is likely to change, and the more the unmarked word order is likely to be affected too. This is confirmed by verb-focus (183a) and subject-focus (183b) in Japanese. In the former, the fp moves to the left of the object. The pitch peak moves too. Since downstep

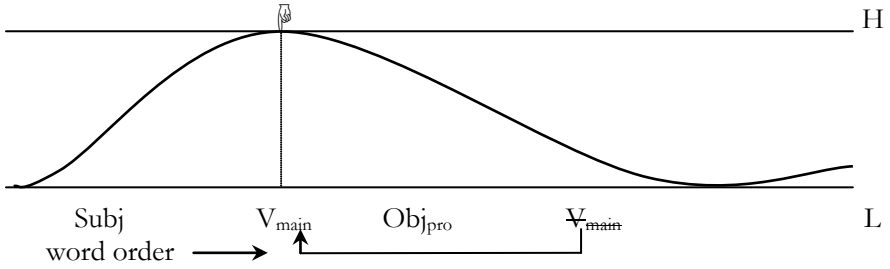
---

(ii) Mari egy KALAPOT nézett ki.
Mari a hat-ACC spotted PRT
'(What did Mari choose?) Mari chose a hat.'
(Szendrői 2003:72-73, (57-58))

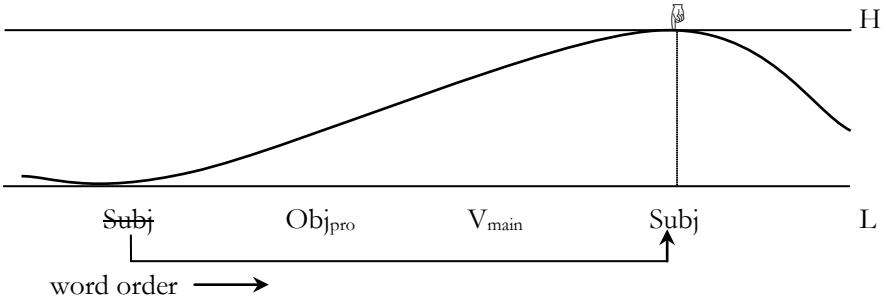
In the proposed system, the fp is located on the main verb in the unmarked case in Hungarian. This claim does not violate principle (174), since broad-focus contains both the main verb and the argument(s). When the fp moves to the left, the unmarked syntactic word order is affected by movement of a focused sentential element, as illustrated in (ii). The pitch peak also moves from the main verb to the focused element to its left.

occurs immediately after the focal point on the main verb, the unmarked intonation pattern changes. The basic word order is also affected by the presence of verb fronting, which results in the SVO order. In the latter case, the fp moves to the right of the object, even across the main verb. The pitch peak occurs on it too. Since downstep does not occur until the subject in sentence-final position, the unmarked intonation pattern changes. The syntactic word order is also affected as illustrated by subject postposing resulting in the OVS order.[137]

(183) a. watashi *tabe-mashi-ta* ano keiki (totteoka-zuni). [Jap.]
I eat-HON-PAST that cake (keep-without)
'I ATE that cake (, not kept it).'



b. keiki-o kat-ta-no-wa *watashi*-desu (, haha-de-naku).
cake-ACC buy-PAST-GEN-TOP I-be mother-be-not
'It's me who bought the cake (, not (my) mother).'

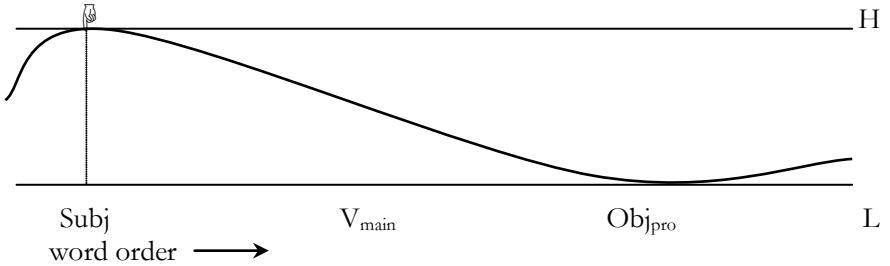


As we have seen so far, the focal point indicated by the fp always coincides with the pitch peak. It is predicted that there are i) cases in which they both move, but the syntactic word order is not affected, and ii) cases in which they do not move, but the syntactic word order is affected. Case i) is illustrated by subject-focus in English (184). The fp and the pitch peak moves to the subject

[137] Subject-focus could also be analyzed in the way that the fp moves to the right and occurs on the subject and the VP containing the object and the verb is fronted. In either way, both the unmarked intonation pattern and the unmarked word order are affected in this construction.

in sentence-initial position. Since downstep occurs immediately after the subject, the unmarked intonation pattern is affected. But the basic word order SVO is not affected.[138]

(184) JOHN (, not MARY,) likes it.

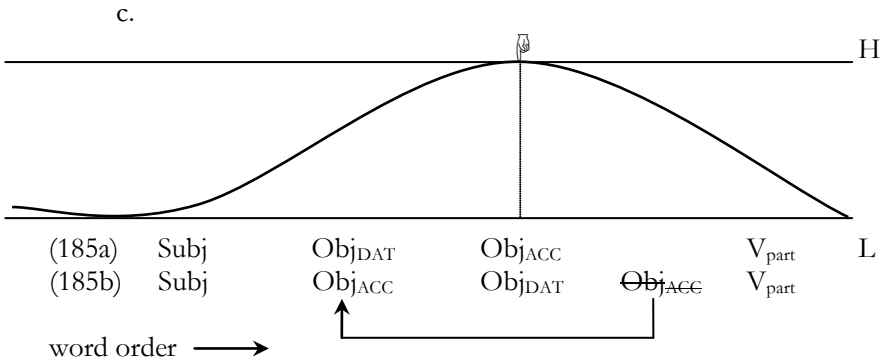


Case ii) is illustrated by scrambling in German (185). (185a) is broad focus, i.e. the unmarked case in which the pitch peak occurs on the sentential element immediately preceding the past participle main verb, i.e. on the direct object *das Buch* 'the book'. (185b) is contrastive argument-focus of the indirect object *dem Kind* 'the child'. The direct object *das Buch* is fronted due to its given status, and the syntactic word order is affected here. As illustrated in (185c), however, the fp occurs on the unmarked position, i.e. on the sentential element immediately preceding the past participle main verb, and the pitch peak also occurs on it in both cases.

(185) a. Hans hat dem Kind das BUCH gegeben. [Ger.]
Hans-NOM has the-DAT child the-ACC book given
'Hans gave the child the book.'

b. Hans hat das Buch dem KIND gegeben.
Hans-NOM has the-ACC book the-DAT child given
'Hans gave the book to the CHILD (, not to her MOTHER).'

[138] In French subject-focus, the unmarked word order is also affected by the use of cleft constructions:
(i) C'est JEAN qui l'aime. [Fre.]
it's Jean who it likes
'It is JEAN who likes it./JEAN likes it.'

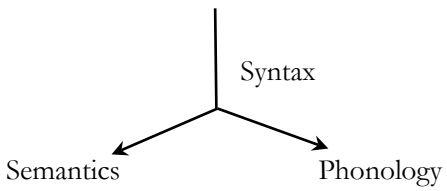


In sum, information structure mediates between grammatical components to express the information flow of a sentence. The basic concept is focus. When the focal point occurs on the unmarked position, the pitch peak occurs on it too; the unmarked intonation pattern does not change. The basic syntactic word order is not affected either. When the focal point moves, however, the focus of a sentence occurs on a marked position. The pitch peak moves too, and downstep starts from the marked position; thus, the unmarked intonation pattern changes. The syntactic word order is also likely to be affected by movement of a sentential element.

5.3. In which grammatical component does Object Shift occur?

In this section, I discuss the grammatical component in which OS can occur. We saw in chapter 2 that generative grammar has traditionally assumed the 'interpretive' model illustrated in (186). According to this model, syntax 'mediates' semantics and phonology. That is, after a syntactic structure is constructed in the syntactic component, it is sent to the semantic component and assigned an interpretation. It is also sent to the phonological component and assigned sound properties.[139]

(186) Traditional 'Interpretive' Model in Generative Grammar:



[139] Recall that in the current phase framework since Chomsky (2000), the syntactic structure is sent to the phonological component at the Spell-Out of each phase in the course of derivation.

We also saw the theoretical change into the phase-cartographic framework. The current phase framework assumes not only that the semantic component is uniform for all languages but also that syntactic operations proceed uniformly for all languages (Chomsky 2001, Berwick and Chomsky 2001). This claim is ensured by the cartographic system (Rizzi 1997, Cinque 1999), in which the position where a sentential constituent is located in syntax, whether as a result of base-generation or movement, must correspond to the interpretation that it receives in the semantic component. In this theoretical framework, a category is interpreted in the moved position.[140]

One of the corollaries of this theoretical framework is that movement cannot occur in the semantic component: the interpretation that a category receives in the semantic component is derived from the fact that it has already moved to and been located in the corresponding structural position in syntax. Thus, the possibility that OS occurs in the semantic component disappears.

Another corollary of the phase-cartographic framework is that optional movement cannot occur in syntax. For movement to occur, a dichotomy between the interpretation that a category receives in situ and the one that it receives in the moved position must be present: the former interpretation must differ from the latter. If such a dichotomy is not observed, movement of that category cannot be regarded as syntactic movement. Such movement has been dealt with as movement in the phonological component, which, by definition, does not affect the change of meaning.[141] We have long seen that OS is optional in many Scandinavian varieties. Thus, one way to account for OS, which is optional and does not affect the change of meaning, within the current phase-cartographic framework would be to say that OS occurs in the phonological component.

A possible derivation, e.g. of Verb Topicalization, (Swe.) *kysst har jag honom inte* 'I haven't kissed him', on this assumption would proceed as follows.[142] The contrastively focused past participle moves to sentence-initial

---

[140] The phase framework and the cartographic system in fact must compensate for each other. On the side of the phase theory, to argue that syntax and semantics are uniform for all languages, the interpretation a category receives must be ensured by the position where it is located in syntax, which is argued in the cartography. On the side of the cartographic system, to argue that a category universally receives the same interpretation in a certain syntactic position, the syntactic computation must be uniform, with a category moving to the same syntactic position to receive a certain interpretation, for all languages, which is argued in the phase theory.

[141] See Chomsky's (2001) argument that movement of a main verb, which can appear either in v* (e.g. English), in T (e.g. French) or in C (e.g. Swedish), though its interpretation does not differ between languages, is a phonological movement.

[142] It is controversial whether main verb movement occurs in syntax or phonology. Chomsky (2001) argues that it occurs in phonology, for the reason stated in footnote 141. Some literature (e.g. Matushansky 2006) argues against that claim, saying that it occurs in syntax. Thus, I take Verb Topicalization as an example here, since the interpretation of contrastive focus of the raised past participle is obvious, and thus it must occur in syntax.

position and the subject and the Aux also move, resulting in *[CP kysst har jag inte honom]*.[143] This structure is spelled out and sent to the phonological component. The pitch peak occurs on the past participle that carries the focus of the sentence. Downstep must occur after it. The object pronoun moves across the negation to cause downstep, resulting in *kysst har jag honom inte*. The pitch remains low on the negation.

It is quite dubious, however, whether and how movement in phonology can be carried out in a theoretically principled way. First, there are no possible candidate features that are directly involved in raising a category in phonology. Syntactic movement requires that a functional head is assigned an EPP feature that triggers movement. As a purely syntactic feature, however, it cannot be directly involved in raising a category in the phonological component. It is not possible to assume that the EPP is assigned to a functional head in the phonological component, since the spelled-out structure cannot be involved in a syntactic operation again.

Assume that in the case of OS, an object pronoun is assigned a phonological feature called [downstep]. This feature could be required by the interface with phonology and cause downstep after the syntactic structure of the OS construction is spelled out. However, that feature itself is irrelevant to movement and does not affect the presence or absence of pronominal shift in phonology.

Secondly, under the traditional interpretive model, it is assumed that semantic/morphophonological realization is derived from some syntactic features present in syntax. For instance, the interpretation as focus in the semantic component is derived from [Foc] present in syntactic component. The morphological realization of agreement in the phonological component is derived from the phi-features present in syntax. In the same way, assume that movement in phonology results from the realization of the feature called [movement in phonology] present in syntax, and that the carrier of this feature moves in the phonological component.

Let us apply the argument above to OS. OS is optional in the Scandinavian languages. Under the assumption of the uniformity in syntax and semantics, it cannot be assumed that the [movement] feature enters a syntactic derivation in some cases but does not enter in others: the syntactic computation would differ between the two cases. Assume that the [movement] feature always enters a derivation, and that it may be realized in some cases (, which results in the presence of pronominal shift) but may not be realized in others (, which results in its absence). It is unclear whether it is justified that the [movement] feature can be allowed not to be realized, since the presence of the EPP always entails movement of a category in syntax.

All in all, it is quite doubtful that movement in phonology can be

[143] In the account here, I leave aside all detailed issues, e.g. Spell-Out of v*P, the assumption that only the complement of a head is spelled out, etc.

derived in a principled way under the current theoretical assumption. For movement in phonology to be feasible in a principled way, a more elaborate derivational mechanism needs to be worked out.[144]

In the thorough discussion of OS from the intonational perspective, we have established that the intonational properties that are not part of the syntactic component under the interpretive model do account for all aspects of OS, i.e. the obligatoriness, optionality and absence of OS. As argued in § 5.1, an object pronoun moves to cause downstep and eliminate a focal effect on the negation on one hand and to maintain the focus of a sentence on the main verb on the other. A possible pitch peak on the negation must be eliminated for the pitch peak to coincide with the focal point on the main verb, as stated in principle (174). Thus, as long as all the *syntactic* behaviors of object pronouns can be accounted for in a principled way in terms of the intonational properties, a different way to account for OS is to say that it occurs in syntax, driven by the intonational properties.[145]

Consider how 'intonation-driven' syntactic movement proceeds by taking Verb Topicalization, (Swe.) *kysst har jag honom inte* 'I haven't kissed him', as an example again. In the same way as in movement in phonology described previously, the following structure will be constructed: *[$_{CP}$ kysst har jag inte honom]*.[146] Since the sentence-initial past participle is contrastively focused, the focal accent and the pitch peak must occur on it. Since the negation is also prominent, the object pronoun moves across the negation to cause downstep on the latter, resulting in *kysst har jag honom inte*. The pitch is low on the negation.

The difference between 'intonation-driven' syntactic movement and movement in phonology described previously is that all operations proceed in syntax. The information that the focal accent and the pitch peak must occur on the past participle enters from phonology in the course of syntactic operations. That information is the trigger of pronominal shift to cause downstep. That some information entering syntax from phonology is the source of pronominal shift indicates that syntax needs to accept some feedback from phonology in the course of syntactic derivations.

Though it is unclear how 'intonation-driven' syntactic movement can be carried out in the current phase-cartographic framework, it is feasible in the system of alignment between syntax, information structure and intonation

---

144 The argument here applies to verb movement too. As stated in footnote 141, verb movement is claimed to be a phonological movement. However, there are no principles/features that determine when and where a main verb moves (or does not move) in phonology. A verbal head could freely choose the position where it moves (or choose not to move). But it appears in either v*, T, or C, depending on languages. It is unclear how this regularity in each language can be accounted for.

145 Or, OS is a 'prosodically-informed syntactic movement', as Vincent van Heuven (p.c.) suggests.

146 Again, I leave aside all detailed issues here.

presented in § 5.2 and, e.g. in the model that assumes the parallel architecture of syntax, semantics and phonology (Jackendoff 2010). In those systems, it is assumed that the grammatical components directly interact with each other. The direction of the interaction is not unilateral, but bilateral. It is possible for the phonological component to directly interact with the syntactic component and affect syntactic operations. I leave the details of how to construct syntactic structures driven by the intonational properties for future research.[147]

[147] See Cheng and Rooryck (2000), who propose to encode intonational properties as a syntactic feature. They observe that the *wh*-in-situ contruction differs from the *wh*-movement construction in that the former has particular intonational properties. They assume that the syntactic feature that is translated into a particular intonation in phonology is included in the selection of the lexical items that enter the derivation of *wh*-in-situ constructions. They propose this system under the *Minimalist Program* (Chomsky 1995), in which the uniformity of syntax and semantics was not assumed.