# How Focus and Givenness Shape Prosody*

*Caroline Féry & Shinichiro Ishihara*

*Abstract*. A model of how syntax and information structure (focus and givenness) shape prosody is proposed which keeps phrasing and tonal effects apart. It is argued that the prosodic effects of syntactic structure and those of information structure should be kept apart. It is shown that in German and Japanese, syntactic structure primarily influences prosodic phrasing, which we assume to be recursive. Information structure, on the other hand, influences tonal structure, keeping phrasing intact. In a comparison between the two languages, it becomes apparent that prosodic domains corresponding to focus and givenness domains are subject to tonal readjustments. A further point made in the paper is that the amount of downstep and reset of register domains is language-dependent.

*Keywords:* focus, givenness, prosody, phrasing, register

## 1. Introduction

This chapter proposes a representation of prosody that distinguishes formally between the imports of syntax, which acts on the formation of prosodic phrases, and those of information structure (focus and givenness), which affects f0 register scaling. It is proposed that the representation of the syntax-prosody interaction has to be able to express both phenomena separately in order to keep their prosodic

* We would like to express our gratitude to Lisa Selkirk, Elsi Kaiser and Malte Zimmermann, whose comments have helped to improve the model presented here. Our work is part of project A1 of the SFB 632 on Information Structure in Potsdam and Berlin, financed by the DFG. Many thanks also to our colleagues Gisbert Fanselow and Ingo Feldhausen.

effects apart. It is also proposed that the resulting prosody is partly language-dependent.

Section 2 introduces the model with German and Japanese, as well as with English, to a lesser extent. Starting with the relationship between syntax and prosodic phrasing, we can lean on a vast amount of literature on the subject, which agrees that prosodic phrasing is mapped from syntax. How exactly this happens and which constituency originates from the mapping have been the subject of extensive discussions, and we will not at this point present an all-ready formalisation of syntax-based phrasing. In section 2.1, we will assume that prosodic phrasing involve recursive constituents.

The second component of our model is the effect of information structure on prosody, discussed in Section 2.2. Focus and givenness, the two information structural properties considered in this paper, change the f0 registers that a speaker uses at a certain point in a sentence. Focus enlarges it and givenness compresses it, and the height of pitch accents and boundary tones are changed accordingly. Crucially, these effects are limited to f0 scaling and do not affect prosodic phrasing, at least in the languages considered here.

Section 3 illustrates the proposal with empirical studies from German and Japanese. Section 4 sums up the main aspects of our proposal in comparison with earlier ones.

## 2. Model

### 2.1 The impact of syntax on prosody

Our main claim is that the prosodic effects of syntactic structure and information structure should be kept apart. For the first part, we assume that syntactic structure

is mapped onto prosody as prosodic phrasing, and that prosodic phrasing is recursive.

In an all-new sentence, that is, in a sentence without new-focus-given partition (see section 2.2. for definitions), the formation of prosodic phrases as well as the tonal pattern and scaling depend entirely on the morpho-syntactic structure. Prosodic phrases have heads in the form of abstract grid positions, which may be realized as pitch accents, as is the case in English and German. The prosodic heads are not necessarily realized with pitch accents, but can be expressed by duration, intensity or even by completely different criteria, like tones or segmental modifications. The prosodic heads correlate with metrical or hierarchical prosodic structure (see for instance Halle & Vergnaud 1980, Selkirk 1984, Nespor & Vogel 1986). An example of such a metrical structure appears in (1) for English. In most studies, relevant levels of phrasing have been assigned a variety of names, like Minor and Major Phrases, phonological phrases and intermediate phrases, accent domains and rhythmic groups, among others (Selkirk 1984, Nespor & Vogel 1986, Ladd 1990, Hayes 1995). However, as soon as the syntax-prosody mapping suggests additional levels of phrasing, such models have to add new names for domains, or allow recursivity of at least some of the levels. In the following, we give up the distinction between smaller and larger domains, and prefer to consider prosodic phrasing as a recursive structure. To keep the terminology as theory-neutral as possible, we call the levels mapped from syntactic constituents 'p-phrases.' Following proposals by Wagner (2005) and Ito & Mester (2006), p-phrases can be embedded into each other. All levels of phrasing shown in (1), except for the lower one, which is the level of the prosodic word (PW) and the upper one, which is the level of the intonation phrase, are p-phrases. It is easy to see that p-phrases can be added when the individual constituents are extended.

(1) ( x ) IP

( x ) ( x ) p

( x ) ( x ) ( x )( x ) p

( x )( x ) ( x )( x ) ( x )( x )( x ) ( x ) PW

Princess Diana's sudden death has been the source of many speculations

In an unmarked case like (1), i.e. in the case of an all-new sentence without any focused or given constituents, the p-phrases of a given level are scaled relative to one another in a completely predictable way (see Kratzer & Selkirk 2007, Truckenbrodt 2002 etc). In the following, we make use of the phonetic abstraction of the phrasal reference line introduced by Van den Berg et al. (1992), a register line of constant height during an intonation phrase (such as a matrix clause), running at the height of the domain-initial peak. To account for downstep among smaller domains, the phrasal reference line is progressively downstepped, as illustrated in Fig.1. In other words, smaller downstep (among accents) is embedded in larger downstep (modeled by the phrasal reference line). Every phrase is downstepped as compared to the preceding one. This downstep takes place at all levels of phrasing. Each prosodic level defines its own downstep pattern, in which the reference top line of every phrase is lower than the reference top line of the preceding phrase of the same category (see also Bruce 1977, Ladd 1990, Truckenbrodt 2002 and Féry & Truckenbrodt 2005 for similar models). In a complex sentence, downstep is thus recursive and a property of embedded prosodic phrases.

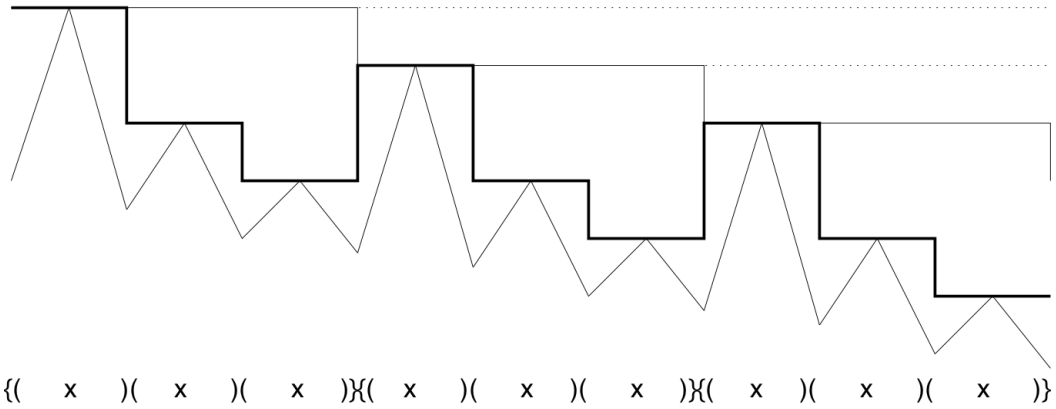
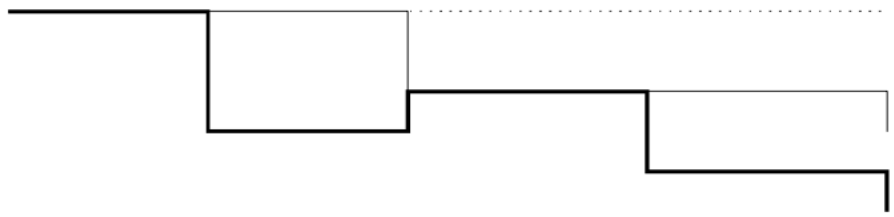



Fig.1: Embedded downstep pattern

Applied to (1), such a model delivers an intricate pattern of downstep relationship between the different phrases at all levels of phrasing, part of which is shown in (2). The highest reference line starts with the first high tone of the intonation phrase and remains available until the end of its domain, here the end of the sentence. In the illustration, it is the dotted grey line. In other words, the top lines do not disappear when following ones become more pervasive. They survive until the end of the relevant domain, and can determine subsequent scaling. This accounts for embedding of register domains inside one another, like the one described below for Japanese.

The immediately lower level of phrasing introduces the first downstep. It separates the subject of the sentence from the VP. The first sister of this downstep relation is at the same height as the beginning of the intonation phrase, but the second part is one step lower. Again, (2) indicates the domain of every reference line with thinner lines. In the first part of this relationship, the grey line stands for the domain of the first sister, and the thin black line for the domain of the second sister. Within these constituents, smaller sister constituents are again inducing a downstep relationship, and so on. Féry & Truckenbrodt (2005) found for German that at a higher level of

prosodic constituency, when clauses are coordinated, a deeper level of embedding induces a steeper downstep.[1]



( x )
( x ) ( x )
( x ) ( x ) ( x )( x )
( x )( x ) ( x )( x ) ( x )( x )( x ) ( x )

(2) Princess Diana's sudden death has been the source of many speculations

F0 registers are calculated relative to the preceding phrases of the same level of prosodic phrasing (see Pierrehumbert 1980, van den Berg et al. 1992, Truckenbrodt 2004, Féry & Truckenbrodt 2005). The most evident cue for registers is the height of initial pitch accents in each prosodic domain, but the boundary tones at the end of prosodic domains can also give indication as to the relevant top lines. The boundary tones can be subject to reset (see Truckenbrodt 2002), and the height reached by a reset boundary tone is determined by the top lines. The calculation of register relationship can be complicated by embedding of register downstep, upstep, final lowering, as well as by purely tonal assimilations and dissimilations (as those found by Féry & Kügler, 2008 for German). But once these effects are factored out, the pattern is straightforward. Though we cannot exclude that some languages do not present a pattern of downstep between prosodic phrases of the same level, at least in

[1] This result is tentatively replicated for smaller prosodic domains in Fig.1, though no empirical

Japanese and in German, downstep seems to be a pervasise phenomenon, which is crucial in the perception of emphasis between accents.

## 2.2. Information structure

In this paper, we are mainly interested in the effect of focus and givenness on the prosodic structure of sentences from a theoretical perspective, Our main claim is that focus and givenness affect the F0-scaling of certain prosodic domains, but do not directly affect prosodic phrasing.

Before illustrating our model, we start the discussion with some definitions of relevant information structural categories, namely focus and givenness. Other information structural categories, like topic, are left out of consideration.

We assume that focus is realized by prosodic prominence, as formulated in (3) (Jackendoff 1972, Truckenbrodt 1995, Büring 2001), though we do not try to define prominence, but assume that it can be realized in different ways (see also Büring, this volume).

(3) Focus Prominence

A focus is realized by prominence in its focus domain.

Throughout the paper, we use the term 'focus' in the sense of Rooth (1985, 1992), an element that singles out referents from a set of alternatives. Focused elements are bound by the focus operator ~. We do not distinguish between what has been called 'narrow focus', 'contrastive focus', and 'identificational focus.' However, we clearly distinguish it from what has been called 'broad focus' or 'information

---

results is available so far that would confirm this hypothesis.

focus', i.e., elements that are new in the discourse. Following Selkirk (2008) and Katz & Selkirk (2007), we assume that discourse-new elements are in the 'default' state, neither bearing foci nor being given.

Prosodic domain of focus (DF) corresponds to the semantic 'scope' of focus, and (3) implies that focus is interpreted and gets its prosodic prominence in this domain. The domain contains the focused phrase and identifies the background information relevant to the semantic denotation of focus (see chapter 4 in Truckenbrodt 1995 for explanation of the focus domain). Focus domain does not necessarily coincide with a single prosodic constituent, as will be illustrated in section 3. In our examples, focus will be formally indicated with a feature F, giving the scope of the focus. As has been shown by Féry & Samek-Lodovici (2006) for embedded foci and by Féry & Ishihara (2009) for Second Occurrence Foci, (3) can be violated when prosodic considerations are overriding the need for a focus to be prominent in its domain.

The other important terminological concept is 'givenness', for which a modified version of Schwarzschild's (1999) definition is adopted. A given constituent is entailed from the context in a precise semantic sense. His proposal amounts to freely assign F-marks instead of letting them percolate along the syntactic tree, like in Selkirk's (1995) proposal. But the two constraints in (4) are restricting the occurrence of these marks. (4a) entails that no F-marking indicates givenness, and (4b) keeps the number of F-marking to a minimum.

(4) a. GIVENness: A constituent that is not F-marked is given.

b. AvoidF: Do not F-Mark.

We depart from his analysis in that instead of assigning F-marks to all non-given elements, we assume that all given elements are G-marked. Féry & Samek-Lodovici

(2006) show that G-marks for givenness are needed, as well. Except when they are focused, given constituents are not accented, a state which is expressed by (5). This constraint directly addresses the prosodic pattern of discourse informed constituents.

(5) Destress-Given

A given phrase is prosodically non-prominent.

The present paper proposes a more precise articulation of the relationship between new/focus/givenness and prosodic structure than (3) and (5) alone are able to achieve. It is proposed that, in German and Japanese, the influence of information structure is mediated through f0 registers corresponding to focus and givenness domains.

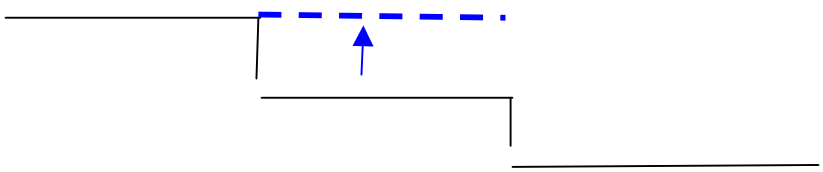
With these definitions, we now illustrate our model. As mentioned above, information structure is reflected in changes in register scaling of prosodic domains and/or focused element.[2] In other words, information structure does not manipulate the boundaries of prosodic phrases as they have been defined by syntax, but instead changes their pitch registers by widening or narrowing them. The most immediate effect of this manipulation is that pitch accents can be higher or lower than in the unmarked situation, according to their focus or given status. If a sentence contains a focus, the F0-register of the focus is affected such that its reference top line is raised, provoking a sudden boosting of the pitch accent correlating with the focused word or exponent. And when given material appears in a sentence, on the other

[2] We are aware that some languages reflect information structures by other prosodic ways, or even by other grammatical components. See for instance, Fiedler et al. (this volume), Hyman (this volume).

hand, this part of the sentence (if any) is compressed.[3] *We assume that the higher pitch accent because of focus is the consequence of the register change within the focus domain, and that lower F0 realization for given material is the result of the compression of the prosodic domain associated with givenness.* The height of the individual pitch accents, generally understood as heads of prosodic phrases, but in some other cases, just the most prominent part of the focus domain, is the result of the transformation of the reference top lines rather than the result of directly boosting or lowering the individual tones associated with accents. One important difference with models directly manipulating the pitch accents is that the relationship between different parts of the sentence is changed. We claim that information structure changes the scaling of the entire sentence, instead of targeting only the most prominent pitch accents. We return to this point below.

When a narrow focus disturbs the unmarked and regular downstep pattern, the top line of the domain corresponding to the focus is raised. This is illustrated in Fig.2.

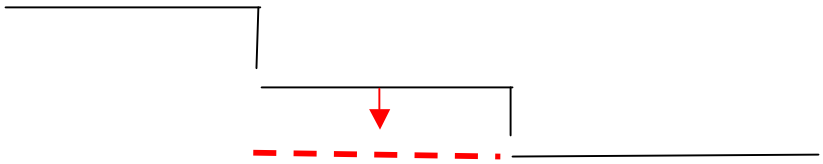
Fig.2. Boosting of the top line of the F-marked.



By contrast, givenness lowers the top line of the given domain, as shown in Fig.3.

[3] As Michael Wagner observed (p.c.) in a sentence ending in a continuation rise, the register can also be narrower because the bottom line is raised.

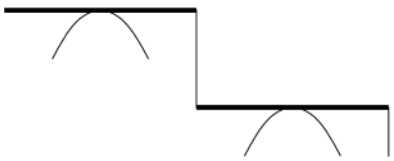
Fig.3. Lowering of the top line of the given domain.



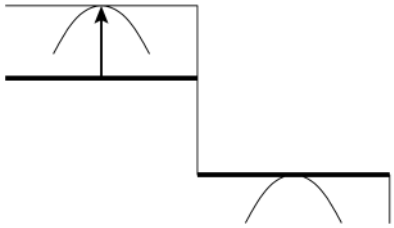
Pitch accents are scaled relative to these top lines. An individual accent will be higher or lower than in the unmarked case, depending on its status as focused or given, because it is constrained by the top line of its domain. Different cases are illustrated in Fig.4. Fig.4a shows the default pattern, in which two prosodic domains are downstepped relative to each other. In Fig.4b, narrow focus raises the top line of the domain in which it appears. In Fig.4c, givenness lowers the top line. Finally in Fig.4d, both effects are present. This can be observed in Second Occurrence Focus (see section 3).

Fig.4 Pitch accents constrained by the top lines of their respective prosodic domains.

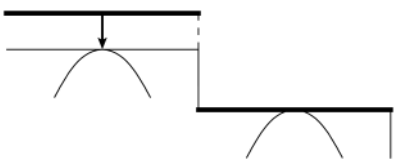
a. Default pattern



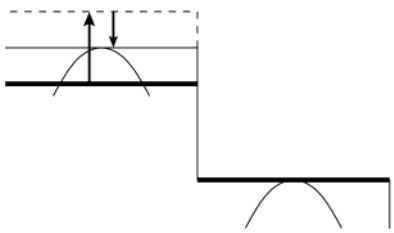
b. Boosting by focus



c. Lowering by givenness

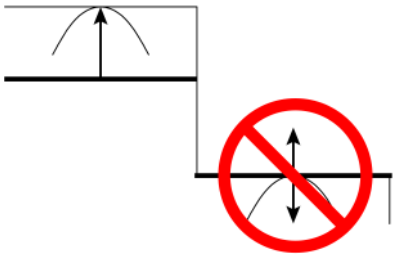


d. Given, focal material



An additional input to the model is the prenuclearity and postnuclearity of the relevant domains. In Japanese, German and English at least, there is a difference between prenuclear and postnuclear effects of givenness. Prenuclearly, i.e. before the last accent of a sentence, the effects of syntax on prosody are pervasive. All prosodic domains are kept more or less intact, even if givenness may lower the relevant top lines. Postnuclearly, however, the top lines are lowered to a minimum, as shown in Fig.5, and are confounded with the baseline. There is no room anymore for the realization of pitch accents. In German, deletion of postnuclear accents is due to the extreme compression or reduction of register. This is especially true when the last accent corresponds to a contrastive narrow focus and when the material following this focus is given. In Japanese, a similar effect can be observed, but to a lesser extent. The compression is not so radical as in German (see Ishihara & Féry, in prep for a comparison between these two languages and Hungarian).

Fig.5. Postnuclear suppression of register



We have assumed until now that prosodic phrases are raised or lowered. It will be shown below that focus and givenness domains do not need to be isomorphic to prosodic domains. Though we have results only for Japanese so far (see below), we assume that the domain of manipulation of reference lines can be smaller or larger than prosodic phrases in German, as well.

To sum up the proposed model, phrase formation is accompanied by prosodic heads, often in the form of pitch accents, whose heights are adjusted to default reference lines. When the sentence has no new-focus-given partition, that is when it is all-new, all prosodic phrases are downstepped relative to each other, and all accents are realized as predicted by the syntax, modulo some variations addressed in the next section. But as soon as information structure comes into play, the reference lines are changed, and pitch accents may become more prominent because the register of their domain is extended. Alternatively, they may become nearly undetectable because their register lines are compressed.

## 3. Empirical results

This section provides empirical evidence for our model. The first set of data, discussed in 3.1, supports the view that only syntax is active in the formation of prosodic phrases. It is in sharp contradiction with models claiming that syntax and information structure have a similar power to shape phrasing (see Gussenhoven, Truckenbrodt). The second set of data, discussed in 3.2, addresses pitch scaling. Our model assumes that pitch accents are scaled to reference lines, which are coextensive with focus domains and givenness domains. It can be contrasted with models taking pitch accents to be directly manipulated. The latter kinds of models do not make any claim as to the scaling of other accents in the sentences, as well as boundary tones which should be blind to information structure. Our model, by contrast, assume that pitch accents and boundary tones, being scaled to reference lines which are in relationship with other reference lines of the same sentence, are radically transformed by a changed information structure.

## 3.1 Prosodic phrasing

It is usually assumed in the literature that, in German, a p-phrase requires a head, which is realized in the form of a phrasal pitch accent, and in Japanese, a p-phrase is the domain of downstep between lexical pitch accents. We show that both criteria are not conclusive, and that a more abstract concept of prosodic phrasing, relying entirely on syntactic structure is to be preferred.

### 3.1.1 Prosodic phrasing in German

A number of researchers link the role of information structure directly or indirectly to the formation of prosodic phrases. It has been proposed a number of times that focus and givenness change the prosodic phrasing by adding or deleting prosodic constituents. A well-known example for Germanic languages is Gussenhoven (1983, 1992), who formulates the Sentence Accent Assignment Rule (SAAR), reproduced in (6), see also Truckenbrodt (2006).

(6) SAAR (Gussenhoven 1992):

If focused, every predicate, argument, and modifier must be accented, with the exception of a predicate that, discounting unfocused constituents, is adjacent to an argument.

As can be seen from the examples in (7), Gussenhoven strictly relates the presence of a pitch accent to the formation of a prosodic phrase, taking the accent as the head of the phrase. A narrow focus may increase or reduce the number of prosodic phrases. In (7b), the predicate (P) is not adjacent to an argument (A). As a result, argument, modifier (M) and verb are phrased individually, whereas in (7a) predicate

and argument are phrased together by virtue of being adjacent. In (7c), the modifier is given, but both argument and verb are focused, and, as a result, argument, modifier and verb are all phrased together.

(7) a. (Any news?)

[Our DOG disappeared]$_F$ [ÁP]

b. (What happened?)

[Our DOG MYSTERIOUSLY DISAPPEARED]$_F$ [Á] [Ḿ] [Ṕ]

c. (Talking about mysteries…)

[Our DOG]$_F$ mysteriously [disappeared]$_F$ [ÁMP]

The generalizations expressed in SAAR are intuitively adequate and go a long way to explain sentence accent assignment. Still, it may be questioned whether the strict link between prosodic phrases and pitch accents is desirable and necessary. We think that it is not. Even though the prosodic structure is mapped to the syntactic structure (see Cinque 1993, Selkirk 1995 Kratzer & Selkirk 2007, Féry & Samek-Lodovici 2006 for different approaches of how this is done), the assumptions that pitch accents are necessarily heads of prosodic phrases, or that every prosodic phrase is necessarily headed by a pitch accent cause problems in certain cases because of the variation observed in pitch accent distribution. Not only sentences with a marked information structure, but also all-new sentences display a certain amount of variation in their pitch accents. Moreover it remains to be demonstrated that prosodic phrases are truly deleted because of deaccenting of accents.

In a production study (Féry & Herbst 2004) it was shown that in the case of a modifier (adverbial or PP) sandwiched between the object and a transitive verb, the verb is often not accented, though the deaccenting of the verb happened less

regularly than when the direct object immediately preceded the verb. The accent properties of all-new sentences like the one in (8) were compared with (9), containing a modifier which had already been introduced in the preceding question, and which was thus considered as given.

(8) Argument–Modifier–Verb (the VP is new)

{Melina is a real entertainer! How did she entertain you this time?}

*Melina hat eine Arie auf der Wanderung gesungen.*

Melina has an aria on the walk sung

'During the walk, Melina sang an aria.'

(9) Argument–Modifier–Verb (the modifier is given)

{I heard that you had a lot of fun on the walk with Melina. What did she do?'}

*Melina hat eine Arie auf der Wanderung gesungen*.

If accents serve as indicators of a prosodic phrase by virtue of being heads, a single accent on the modifier or on the preceding object correlates with a single prosodic phrase, whereas two accents, one on the verb and one on the constituent preceding it, reveal the presence of two prosodic phrases.

In a comparison between (8) and (9), displayed in Fig.6, it is conspicuous that the modifier was accented more often when it was part of an all-new sentence (8) than when it was given (9), namely in (91% vs 58% of all cases, respectively). This can be seen by comparing the middle bars in Fig.6. Again, the argument was clearly accented in nearly all cases (see the leftmost bars).

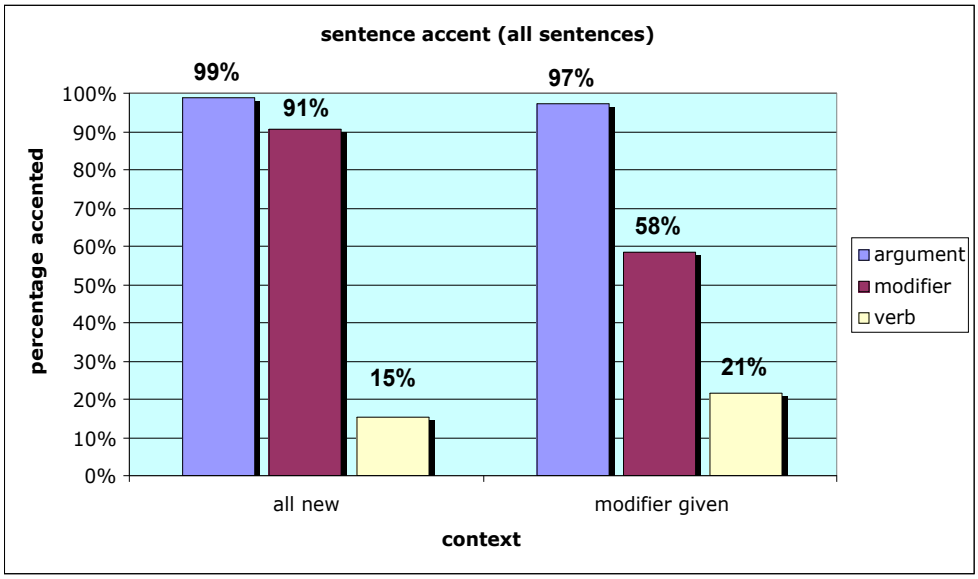



Fig. 6: Percentage of accented phrases (from Féry & Herbst 2004)

A striking discrepancy between the predictions of SAAR and the results obtained in the experiment concerns the accenting of the verb. The percentage of accented verbs does not change much across the conditions. The verb was accented in 15% of the cases when the modifier is new, and in 21% of the cases when the modifier is given, a result going into the opposite direction from the one posited by Gussenhoven. We conclude that there is no correlation between accenting of the verb and accenting of the modifier.[4]

On the basis of these results, it is reasonable to assume that pitch accent assignment is an indicator of phrasing only up to a certain point. The fact that the object was (nearly) always accented confirms the assumption of a metrical grid in which the object is the head of the sentence. But the optionality of the pitch accents on the verb and on the modifier leads us to assume that the verb is part of the prosodic phrase of the preceding constituent, argument or modifier, and that its variable

[4] In VPs consisting in just a modifier plus a verb (and no object), the verb is accented in the majority of cases (unpublished data). In no more than 31 sentences out of 210 (15%), the modifier is the only accented element. SAAR does not predict such a change in the accent structure just because the object is absent.

accenting is the result of other constraints than prosodic phrasing, like different information content of the verb and/or of the modifier for instance.

A recursive phrasing of such example, as was proposed in Féry & Herbst (2004) is to be preferred to a structure predicted by SAAR. The p-phrase of the modifer is embedded in the p-phrase projected by the whole VP and comprising the object and the verb. This is illustrated in (10).

(10) Melina hat [eine Arie [auf der Wanderung]$_P$ gesungen]$_P$

The object is the head of the larger p-phrase, and is assigned a pitch accent because of this property. But both the verb and the modifier can also get a pitch accent, the modifier because it is also a head of a p-phrase, albeit one of a lower, embedded kind, and the verb because it is separated from its head, and also form a prosodic unit of some kind.

The upshot is that an account of phrasing which relies entirely on the physical presence of pitch accents is fragile when it comes to variation. Moreover, it assumes that prosodic phrasing is contingent on phonetic realization, and that each uttered sentence comes with its own prosodic phrasing. The alternative proposed here is that prosodic phrasing is part of the syntactic derivation. It is an abstract structure submitted to phonological and phonetic realization. As such, it is not yet realized, and can not have physical pitch accents. It determines potential locations of accent realizations, which are represented as grid positions, but the actual realization of pitch accents depend on a number of factors, like adjacency to other accents, givenness, and so on.

3.1.2 Prosodic Phrasing in Japanese

In Japanese, pitch accent is part of lexical information and can thus not be taken as defining characteristic of p-phrases in the same way as in German. Instead of pitch accents, tonal scaling (downstep, reset) is used as evidence for the existence of p-phrases. It has long been assumed that p-phrase (i.e., what has been called *Major Phrase* or *intermediate phrase* in the literature) is the domain of downstep (Poser 1984, Pierrehumbert & Beckman 1988, Kubozono 1993, among many others).[5] That is, if a sequence of two pitch accents is subject to downstep, they belong to the same p-phrase. If, by contrast, the second pitch accent is reset to the height of the first one, there is a p-phrase boundary between them. This definition of p-phrase, together with some standard assumptions, however, raises several problems. In this section, we discuss some of them.

One of the widely accepted analyses of p-phrasing in Japanese is Selkirk & Tateishi's (1991) end-based model. They claim that the left edge of a syntactic maximal projection (XP) corresponds to the left edge of a p-phrase. In this analysis, the existence or absence of XP-boundaries is the deciding factor for the existence or absence of p-phrase boundaries, and for the corresponding downstep effect. The analysis predicts that in a left-branching structure like (11a), both N2 and N3 show downstep, while in a right-branching structure like (11b), only N3 does.

(11) a. Left-branching structure: [[ N1 N2 ] N3 ]
Major Phase structure: ( N1 N2 N3 )

b. Right-branching structure: [ N1 [ N2 N3 ]]
Major Phase structure: ( N1 )( N2 N3 )

This prediction is not completely borne out. Kubozono's (1993) experimental data show that N2 in (11b) shows downstep as well, even though the amount of lowering is significantly smaller than in (11a).[6,7] Furthermore, recent studies show that downstep is not limited within what the earlier analyses claimed to be p-phrases. Kubozono (2006) shows that the downstep can be observed after a syntactic XP-boundary, which is considered to correspond to a p-phrase boundary. Experimental results reported by Ishihara (in preparation) also confirm that pitch reset effect at the XP-boundary is only partial, and a smaller amount of downstep effect remains after the XP-boundary. If so, we need to reconsider the definition of either p-phrase or the domain of downstep.

In our analysis, we adopt recursive p-phrasing. By allowing p-phrase embedding, we can not only maintain the assumption that p-phrase is the domain of downstep, but also explain the smaller amount of downstep across XP-boundaries. First, we assume that XPs in the left- and right-branching structure are mapped as embedding

---

[5] We no longer use these terms because once we adopt recursive p-phrasing, as proposed in Ito & Mester (2006), we no longer need the distinction between major and minor phrase, or prosodic phrase and intermediate phrase.

[6] For the sake of fairness, it should be mentioned that in Selkirk & Tateishi's (1991) data, there was no downstep on N2 in a structure like (b), just as predicted by their analysis. We believe, however, that their results are influenced by an (unwanted) effect of focus. The sentences used in their experiment, e.g. the one in (i), are structurally ambiguous between a left-branching and a right-branching parse:

(i) a. Left-branching:
[[Aoyama-no Yamaguchi-no] aniyome-ga] inai
-GEN -GEN sister-in-law-NOM not.there
'We cannot find the sister-in-law of Yamaguchi from Aoyama.'

b. Right-branching:
[Aoyama-no [Yamaguchi-no aniyome-ga]] inai
'We cannot find Yamaguchi's sister-in-law from Aoyama.'

(Selkirk & Tateishi 1991:523)

It is highly plausible that the speakers, who read both sentences and hence were fully aware of the difference in syntactic structure as well as meaning, purposefully tried to disambiguate the two sentences by placing a focus on N2 in the (b)-sentence. If that is the case, Selkirk & Tateishi's results do not show the pure effect of p-phrasing.

[7] Kubozono (1993) tries to account for the difference by proposing a phonetic boosting effect called *metrical boost*, which applies to the phrase on the left side of XP (i.e., where Selkirk & Tateishi's analysis predicts there to be a p-phrase boundary). The crucial difference between the two analyses is that the metrical boost is a phonetic effect of expanding the pitch range on the XP-leftmost phrase, and can apply multiple times, whereas p-phrase boundary insertion is a phonological operation that can be applied only once.

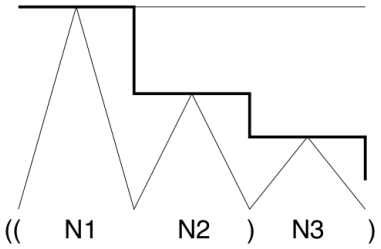
p-phrases, as shown in (12). In the left branching-structure, N1 and N2 form a single p-phrase, which forms a larger p-phrase with N3, while in the right-branching structure, N1 forms a large p-phrase with a smaller p-phrase containing N2 and N3. Second, maintaining the standard assumption that p-phrase is the domain of downstep, we expect a pitch reset at each left-edge of p-phrases.

In the left-branching structure in (12a), all left edges are aligned at the beginning of the entire phrase. Therefore no reset is expected within the phrase. There are two p-phrases, the smaller p-phrase containing N1 and N2, and the bigger one containing the smaller p-phrase and N3. Inside the former, downstep is expected on N2, and in the latter, on N3. As a result, we observe a successive downstep.

In the right-branching structure in (12b), there are also two p-phrases. The smaller one containing N2 and N3, and the bigger one containing N1 and the smaller p-phrase. In this case, we expect a pitch reset at N2, because a left edge of the smaller p-phrase appears here. However, we also expect a downstep effect at this position, because downstep is expected at the larger p-phrase, between N2 and the smaller p-phrase. As a result, we observe a small amount of downstep effect at N2.

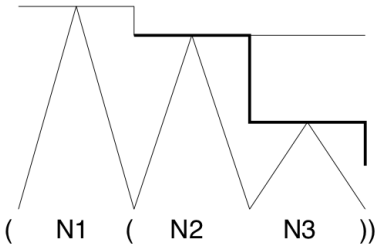
(12) a. Left-branching structure: [[ N1 N2 ] N3 ]

P-phrase structure:




b. Right-branching structure: [ N1 [ N2 N3 ]]

P-phrase structure:




Japanese p-phrase further supports the view that p-phrase is formed recursively, according to syntax. Branching structure is reflected onto p-phrasing, and the different amount of downstep observed in the left- and right-branching structure can be explained by recursive p-structure.

In German, pitch accent distribution does not speak for a one-to-one mapping between accents and p-phrases. A recursive structure, allowing variable pitch accent assignment makes better prediction.

### 3.2 Pitch range and pitch scaling

In this section, results on pitch scaling are summed up which show how tones are scaled differently as a result of varying information structure. As before, results are first shown for German, and in a second step for Japanese.

3.2.1 Pitch scaling in German

Two experiments on German are summarized, one on Second Occurrence Focus (Féry & Ishihara, 2009), and another one on the scaling of pitch accents in sentences with a simple syntactic structure under varying information structure conditions (Féry & Kügler 2008).

In Féry & Ishihara (2009), it is shown that the pitch height of a word depends on its status as First Occurrence Focus (FOF, a focused expression that appears in the discourse for the first time), Second Occurrence Focus (SOF, a focused expression that has already appeared in the discourse and is repeated) or just given (Non-Focus). SOF is illustrated in (13) and (14) with one of the sentences from the object set used in our experiment.[8] The study was a follow-up on studies by Rooth (2004), Bartels (2004) and Beaver et al. (2007), who looked at SOF in postnuclear position in English. In our study, prenuclear position was also considered, which led to clearer results concerning the realization of pitch accents (see also Rooth, this volume, for relevant discussion).

(13) Second occurrence focus: prenuclear

{Many women have invited several relatives to the village fair.

FOF: Aber Eva hat nur ihren Bruder eingeladen.

'But Eva has invited only her brother.'}

SOF: *Nur ihren Bruder hat auch Maria eingeladen.*

'Only her brother.ACC has invited also Maria.'

(14) Second occurrence focus: postnuclear

{Many women have invited several relatives to the village fair.

[8] The experimental sentences were divided into two sets: in one of them the relevant expressions were the subject of the sentence, and in the other one, they were the object.

FOF: Aber Eva hat nur ihren Bruder eingeladen.

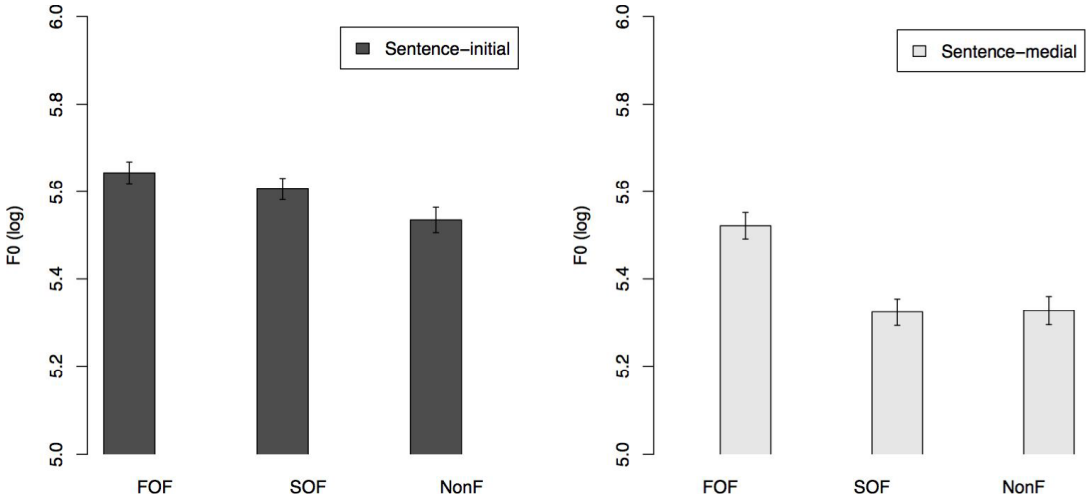
'But Eva has invited only her brother.'}

SOF: *Auch Maria hat nur ihren Bruder eingeladen.*

'Also Maria has invited only her brother.'

As far as pitch accents are concerned, we showed that there is a hierarchy in terms of pitch height among FOF, SOF and non-focus: a FOF is realized higher than a SOF, which in turn is higher than a given constituent (or non-focus). This hierarchy is best observed in the sentence initial position (dark bars in Fig.7, corresponding to the subject set).

In the postnuclear position, pitch accent differences were cancelled between SOF and non-focused, since a postnuclear deaccenting had taken place in this position, and suppressed any potential differences. This can be seen in comparing the final two light bars which stand for the object set.




a. Sentence-initial (Subject set)

b. Sentence-medial (Object set)

Fig.7: Mean log-transformed F0 for FOF/SOF/Non-Focus in sentence-initial/medial conditions (with 95% Confidence Interval)

We analyzed SOF as being simultaneously given and focused, and posited that the intermediate height of a SOF is the reflex of this double status. We also found a difference in duration, reproducing Beaver et al's results. More precisely, a focused expression, be it FOF or SOF is always longer in duration than a non-focused one. The pitch range is thus sensitive to information structure, and assignment of accents is not just a yes or no matter.

The second relevant study for German has been reported in Féry & Kügler (2008) who investigated the production of sentences of the following type.

(15) Subject **Verb** (focus on the verb)

{The animals don't like to fight. Why are they angry with the sheep?}

Weil der Hammel [<u>angefangen hat</u>]$_F$

'Because the sheep started (a fight).'

(16) **<u>Subject Object Verb</u>** (all-new condition)

{'Why were the animals happy?'}

[Weil der Hummer den Rammler eingeladen hat]$_F$

'Because the lobster has invited the buck'

(17) Subject Object **<u>Indirect Object</u>** Verb (focus on the dative complement)

{The buck wanted to introduce the sheep to the lion. Why didn't he do that?}

Weil der Hammel den Rammler [dem Hummer]$_F$ vorgestellt hat.

'Because the sheep introduced the buck to the lobster.'

Three parameters were systematically varied. First, the number of arguments, which varied between one and three (subject, direct object and indirect object). Second, the word order, and third the given-new status of the arguments and the verb. Altogether 26 conditions were created. In total 2340 (130 sentences × 18 speakers) were recorded, and 2277 were analyzed.

The aim of this experiment was to quantify the downstep coming from the syntax, the tone-boosting effect of focus, and the tone-lowering effect of givenness.

Fig.8 shows how downstep regulates the high tones of each argument and the verb in the four sentence lengths.[9] This pattern is thus the result of the syntactic structure which predicts a lowering on each high tone in an intonation phrase.

[9] It must be observed that only 34% of the all-new sentences had this pattern. The other realizations had upstep on the preverbal arguments, or the verb was unaccented.

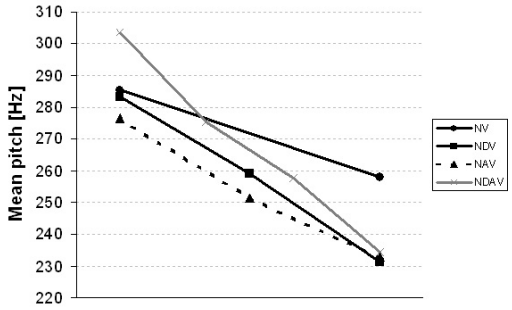
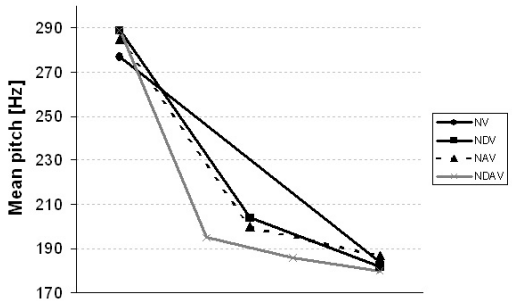
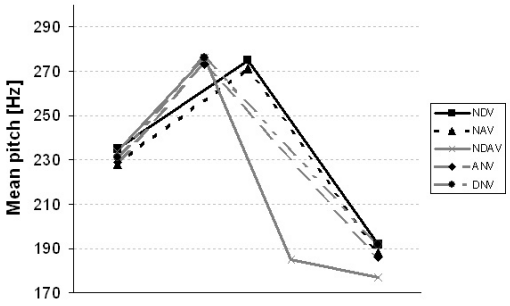



Fig.8 Downstep pattern in all-new sentences

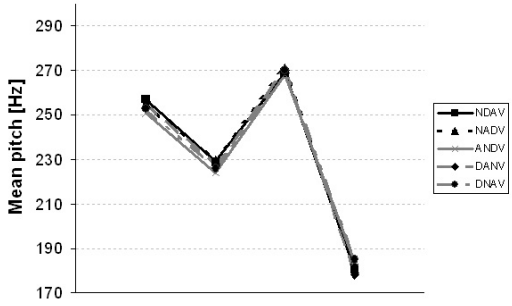
The effect of focus and givenness appears in Fig.9 for sentences with narrow focus on the first (a), the second (b), the third argument (c) and the verb (d). it is easy to see that the focused argument causes an f0 raising of the high tone, and that the other given arguments are lower than in the default case shown in Fig.8.



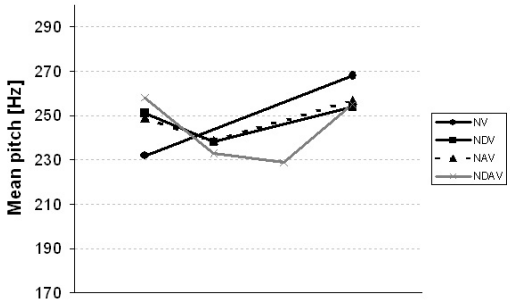

(a) narrow focus on the first argument




(b) narrow focus on the second argument




(c) narrow focus on the third argument




(d) narrow focus on the verb

Fig.9. Representation of pitch peak realization of narrow focus on different constituents.

This experiment confirmed that the downstep and lowering effects caused by syntax and by information structure are distinct. There was no indication that the prosodic phrasing could be distinguished in the same way.

Tonal scaling is thus remarkably sensitive to changes in information structure, and varies with the height of the accented constituents in a very fine-grained way. Postulating only prosodic phrases for these results is not sufficient. When the status of one constituent is changed to focused, SOF or given, not only the f0 height of its accent is affected, but also the height of the other constituents in the same sentence. There is no indication that the relationship between syntax and prosodic structure is changed as well, and even if we wanted to relate the changes in f0 height to prosodic domains, it would be a delicate matter to decide which prosodic level is affected by being neighbour of a focus, and to distinguish between pre- or post-focality. The solution we offer to separate prosodic phrasing and pitch scaling, and treat them as two independent factors entering the prosodic structure of a sentence, can account for these results.

### 3.2 Tonal scaling in Japanese

In the case of Japanese, as well, phonetic effects of focus have often been explained in terms of prosodic phrasing. It has been claimed that focus influences p-phrase structure. Pierrehumbert & Beckman (1988) proposed that focus creates a p-phrase boundary on its left, hence blocking downstep effects. Also, the f0-downtrend observed on post-focal material has often been analyzed as downstep, as a result of

the elimination of all p-phrases after focus (Nagahara 1994).[10] Under this line of analysis, focus directly affects prosodic phrasing, by adding or deleting p-phrase boundaries.[11]

In Ishihara (2007), it is claimed on the base of several experimental results that p-phrase and 'Focus Intonation' (FI) domains are independent from each other. The former is purely syntax-based, while the latter is created by changing the pitch register of the focused phrase (by f0-boosting) and the post-focal material (by f0-compression). The experimental results show that downstep, which lowers the pitch realization of non-initial pitch accents within a p-phrase, and post-focal compression, which lowers the pitch realization of post-focal material, are independent phenomena. That is, a pitch reset after downstep at the following p-phrase boundary, and a pitch reset after an FI can be observed separately. In (18), where there is no focus, the embedded clause verb *nonda* 'drank' and the preceding PP *nomiya-de* 'at the bar' form a single p-phrase. Accordingly, downstep is observed on the verb, followed a pitch reset on the matrix adverbial phrase *imademo* 'still'.

(18) *Declarative sentence (Default pitch contour)*

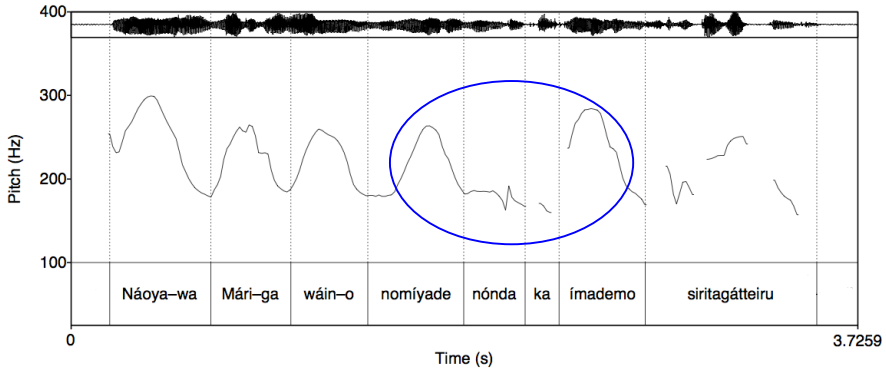
Náoya-wa [$_{CP}$ Mári-ga wain-o$_i$ [$_{VP}$ nomíya-de t$_i$ nónda] to] ímademo omótteru

N.-TOP M.-NOM wine-ACC bar-LOC drank that still think

'Naoya still thought that Mari drank something at the bar.'

[10] See, however, Shinya (1999) who argues against this line of analysis, and supports Poser's (1986) claim that the f0-boosting effect of focus is independent of MaP boundaries. See also Sugahara (2003), who experimentally showed that MaP boundaries can be observed in the post-focal area when it is discourse new.

[11] Although some authors claim that the focus creates an Intonation Phrase prominence (Truckenbrodt 1995, Selkirk 2006a), the basic concept remains the same in that focus modifies prosodic phrasing. See sections 2 and 4 for relevant discussion.



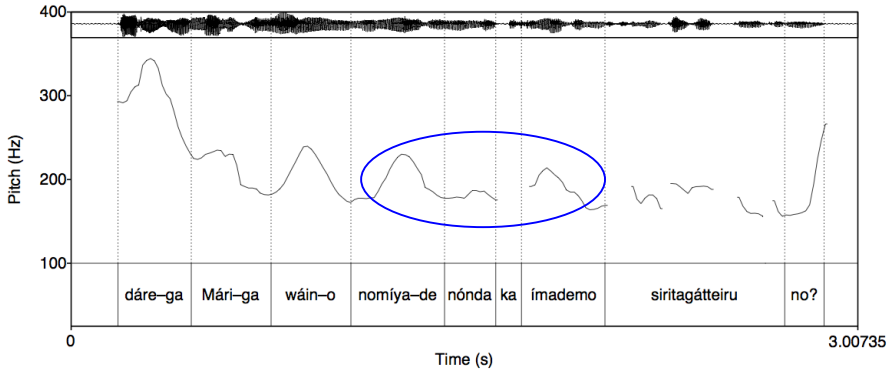

In a *wh*-question like (19), where the wh-phrase behaves prosodically as a focused phrase and hence triggers a FI, the $F_0$-peak on the *wh*-phrase is raised, and the post-focal material is compressed.

(19) *Wh-question (FI starting from the wh-phrase)*

dáre-ga [$_{CP}$ Mári-ga wáin-o$_i$ [$_{VP}$ nomíya-de t$_i$ nónda] to] ímademo omótteru no?

who-NOM M.-NOM wine-ACC bar-LOC drank that still think Q

'Who still thinks that Mari drank wine at the bar?'




It should be noted here that we can still observe a downstep on the embedded verb *nonda*, which is realized lower than the preceding PP *nomiya-de*, and the subsequent pitch reset on the matrix adverbial phrase, which is realized roughly as high as *nomiya-de*, even though the entire pitch contour is compressed due to the

focus in sentence-initial position (see also Fig.10). If we took the FI to be a large p-phrase created by deletion of p-phrase boundaries, as the earlier accounts claim, then we would expect downstep on all the phrases in the sentence, and no pitch reset would be expected. What we see here instead is that the p-phrase structure is maintained within the (more or less evenly compressed) post-focal domain. The normalized mean $F_0$-peaks and valleys given in Fig.10[12] confirm this observation: within the compressed pitch contour in the wh-question (condition B, solid line), the $F_0$-peak on the embedded verb shows a larger fall due to downstep, and that of the matrix adverb shows a pitch reset.

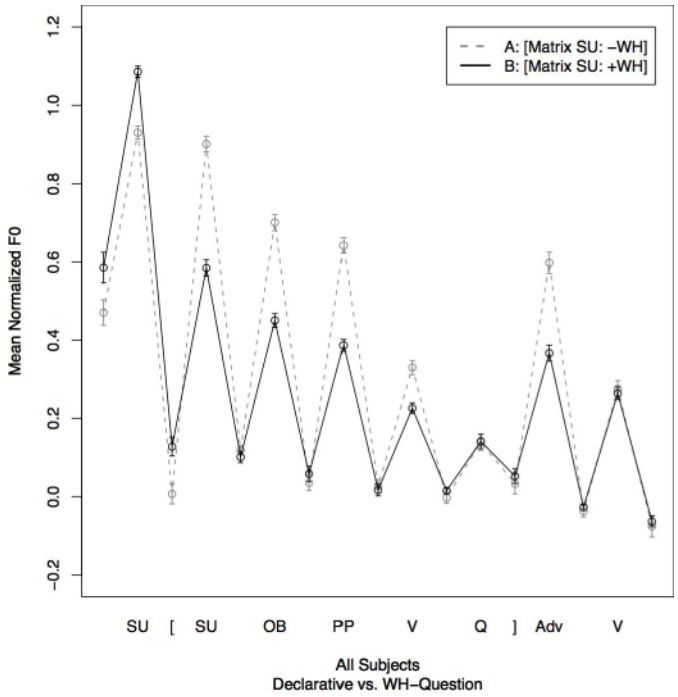



Fig.10: Normalized mean F0 peaks and valleys of (15) and (16)

Furthermore, another experiment by Ishihara (in preparation) reveals clear phonetic differences between focal $F_0$-boosting and p-phrase boundary. In the example below,

[12] The results from 11 subjects are normalized to factor out the individual difference in pitch range. Each condition has 198 samples (6 sentences recorded 3 times per subject, produced by 11 subjects).

the pitch realizations of the first three nouns (N1–N3) were measured. In (20a), there is neither focus nor p-phrase boundary on N3, whereas there is a focus in (20b) and a p-phrase boundary in (20c).

(20) a. *No focus, no p-phrase boundary on N3*

[$_{VP}$ [$_{DP}$ Náoya-no áni-no wáin-o] waingúrasu-de nónda]

Naoya-GEN brother-GEN wine-ACC wineglass-with drank

'(I) drank Naoya's brother's wine with a wineglass.'

b. *Focus on N3*

[$_{VP}$ [$_{DP}$ Náoya-no áni-no náni-o] waingúrasu-de nónda] no?

Naoya-GEN brother-GEN what-ACC wineglass-with drank Q

'(Lit.) Naoya's brother's what did you drink with a wineglass?'

c. *P-phrase boundary before N3*

[$_{DP}$ Náoya-no áni-ga ] [$_{VP}$ wáin-o waingúrasu-de nónda]

Naoya-GEN brother-NOM wine-ACC wineglass-with drank

'Naoya's brother drank wine with a wineglass.'

The results are summarized in Fig.11. When N3 bears a focus, as shown in the solid line in the left graph, the $F_0$-peak on N3 is raised compared to a non-focus. Nothing else is changed. When there is a p-phrase boundary in front of N3, not only the $F_0$-peak on N3 is raised, showing a pitch reset, but also the $F_0$-peak on N2 is lowered. This suggests that focus only affects the $F_0$-peak of the focused phrase, an insertion of a p-phrase boundary not only affects the p-phrase-initial phrase but also the final phrase of the preceding p-phrase. If there is such a clear difference between the phonetic effects of focus and those of p-phrase boundary, they should be treated differently in the phonological analysis.

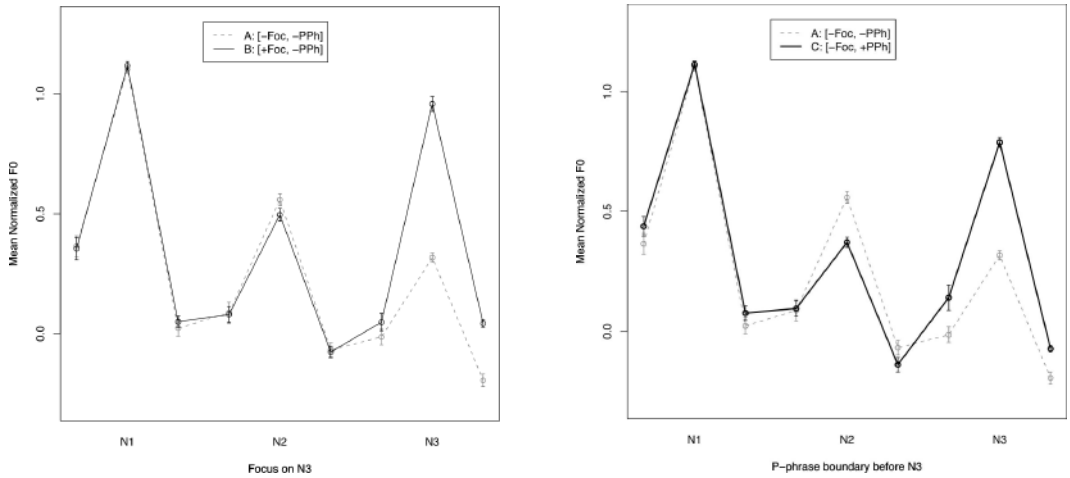



Fig.11: Normalized mean $F_0$ of (19a) (dotted line); (19b) (left); and (19c) (right)

In sum, we find an extreme sensitivity of pitch scaling to information structure, both in German and in Japanese, and in different domains of information structure. It does not seem desirable, and even possible, to explain all effects by changes in the prosodic phrasing.

## 4. Comparison with earlier approaches

In this section, we compare our approach with some of the other models that have been proposed in the literature. The core of our approach is the clear separation of prosodic influence of syntactic structure and those of information structure. This property sharply contrasts with those in many analyses that attempt to capture those effect in a single phonological representation.

Earlier analyses may be divided into two subgroups depending on whether they adopt prosodic phrasing or not. In one group of analyses, information structural feature is assigned directly to morpho-syntactic elements, and both syntactic and information structural effects are encoded directly into prosody (most often as pitch accents), without any direct reference to prosodic phrasing. In another group of analysis, syntactic and information structural features are encoded onto prosodic phrasing.

### 4.1 Pitch accents without mediation of prosodic phrasing

The first type of models which does not make a clear distinction between the effects of syntax and those of information structure posits that pitch accents are directly assigned to morpho-syntactic constituents or to focused constituents. The distribution of default pitch accents is driven by left- or rightmost Nuclear Stress Rules or by projection rules which take the predicate-argument structure of the sentence into account. Pitch accents are usually represented in the phonological representation as positions in metrical grids or metrical trees. In a second step, information structure modifies the default or unmarked pitch accent distribution. Focus constituents gets an accent and given constituents loose their default accent (Jackendoff 1972, Schmerling 1976, Ladd 1980, von Stechow & Uhmann 1986, Cinque 1993, Rooth 1985, 1992, Schwarzschild 1999, Steedman 2000, Büring 2006, Beaver & al 2007 among many others)

A famous representative of this group is Selkirk's (1995) *Focus Projection Theory* and its variants and descendents. F-marked constituents are assigned a pitch accent, and the F-feature percolates up in the syntactic structure.

In Selkirk's analysis, givenness is derived from lack of F-marking (see also Schwarzschild 1999 and Büring 2006). As we saw above, however, such direct

association of givenness and deaccentuation causes problem in certain cases such as second occurrence focus, where given element may and may not bear pitch accent depending on its relative location to the first occurrence focus. In our model, on the other hand, syntactic effects are mapped to prosodic phrasing while the information structural effects are encoded onto pitch register. This allows interaction of the two different effects.

A serious drawback of such an approach is that pitch accents are always modified individually. A focus just changes one pitch accent, for instance, and leaves the remaining of the representation untouched. In other words, no global effect in the scaling of accents can be accounted for by such approaches. If all pitch accents of a sentence have to be modified, this can only be done by addressing every tone one by one. Furthermore, boundary tones or phrasal tones are either completely left out of such representations, or assumed to be anchored in a different part of the phonology, namely in the intonational phonology (Pierrehumbert 1980).

### 4.2 Pitch accents are mediated by prosodic phrasing

In the second group of analyses, prosodic phrasing mediates the syntactic and information structural effects (Beckman & Pierrehumbert 1986, Gussenhoven 1983, 1992, Selkirk 1986, Truckenbrodt 2006).

There are two main assumptions in this line of analysis. The first general assumption is that *all* prosodic effects are explained in terms of prosodic phrasing. This means that prosodic effects of syntactic structure and those of information structure are both encoded onto prosodic phrasing. The second prevalent assumption (at least for intonation languages such as English and German) that comes with this analysis is that heads of prosodic phrases are marked by pitch

accents. With the combination of the two assumptions, syntactic and information structural effects are explained in terms of insertion or deletion of prosodic boundaries.[13]

According to some version of syntax-prosody mapping principle (e.g., Selkirk 1986, Truckenbrodt 1995), syntax determines the location of prosodic boundaries. Once the prosodic structure is determined according to syntax, the head of each prosodic phrase bears a pitch accent, under the second assumption mentioned above. Information structure (focus and givenness), on the other hand, affects the location of pitch accent. Focus requires a pitch accent, and givenness the lack of it. Again, under the second assumption, this amounts to saying that when there is a focus, it must bear a pitch accent, and consequently, become the head of a prosodic phrasing. This forces the modification of prosodic phrasing. A pitch accent assigned to focus forces insertion of a prosodic boundary. The opposite phenomenon takes place for given material. Givenness requires lack of pitch accent, and hence, may eliminate pitch accents assigned according to the syntax-prosody mapping. Under the second assumption, elimination of pitch accent directly means the elimination of prosodic phrase boundary, because no prosodic phrase may exist without its head. All in all, both syntax and information structure affect the realization of prosodic phrasing, and nothing else. A similar view explains earlier approaches to Japanese: an additional upstep associated with focus forces the insertion of a prosodic boundary, and compression of f0 means that prosodic phrases have been deleted.

But such an analysis cannot explain some of the data discussed in this chapter. Japanese data presented in Fig.11, for example, clearly show that prosodic boundary inserted at the syntactic boundary and the F0-boosting triggered by focus behave

[13] Selkirk (2006a) proposed for Bengali that a morphemic tone $[H]_{FOC}$ necessarily aligns with the end of a syntactic focus, creating in this way a prosodic boundary. In this approach, a tone associated with focus creates a prosodic phrasing, and deletion of post-focal prosodic phrases follows.

differently. We need a more elaborated model than insertion and deletion of prosodic phrases, in order to explain the information structural effects on prosody.[14]

### 4.3 Comparison with our model

Our model integrates several main components of the tradition from which it originates. An crucial ingredient is that prosodic phrases are mapped from the syntax. In the case of German and English, the location of a default pitch accent correlates with the location of heads of p-phrases. In the case of Japanese each p-phrase serves as a domain of downstep. The distribution of pitch accents is thus regulated by constraints on the syntax and on the accent projection, aspects which we did not evoke in the paper. In the case of Japanese, each p-phrase serves as the domain of downstep.

Our model differs from the earlier accounts just mentioned in two respects: (i) it does not assume the Strict Layer Hypothesis and adopts a recursive p-phrasing instead; (ii) it denies that a unique prosodic structure is shaped in the same way by syntax and by information structure, and instead, proposes that the prosody responds to different influences in different ways. In particular, we claim that the main import of information structure on prosody is its effect on f0-scaling. As shown in detail in section 2, raising, lowering and deletion of pitch accents are mediated by the effect of information structure on focus and givenness domains. The pitch accents are scaled to top and bottom lines of f0 register of these domains, which can be isomorphic to prosodic domains, but do not have to. These registers are calculated as a function of what precedes and follows. It is this aspect of the model which differs most from the proposals presented above. Many of them

[14] However, it has been shown for diferent languages that prosodic phrasing may be sufficient as an indication of focus. Kanerva (1990) shows for Chichewa that a change in phrasing triggered by focus is not necessarily accompanied by prosodic prominence in the form of pitch accents. The focus structure is thus expressed by the phrasing and nothing else (see also Fiedler et al. this volume).

consider the changes in f0 as the phonetic correlate of a change in the prosodic phrasing. In another line of thought, an accent standing for a narrow focus may be the correlate of a metrical gird position assigned at a different level of phrasing from an accent standing for an information focus (Selkirk 2002). An important consequence of our proposal is that it can account for the global effect that a focus has on the remainder of the sentence. The scaling of all pitch accent is the perceived consequence of the changed structure of the registers as a consequence of the focus-new-given partition. It must not be postulated that every pitch accent is subject to a special rule principle or constraint which raises or decreases it.

## 5. Conclusion

In this paper, we have presented both empirical results and theoretical reflections on the interaction between syntax, information structure and prosody. It has been shown that the effects coming from the syntax and those coming from the information structure should be kept apart in order to understand better how the prosodic structure is obtained at the level of the sentence. We have compared an intonation language with free assignment of pitch accents (German) with a pitch accent language in which pitch accent distribution is lexically governed (Japanese). Syntax regulates prosodic phrasing and default pitch scaling, whereas information structure acts on the f0 of entire focus and givenness domains. A side aspect of our proposal is the systematic recursivity of prosodic phrases.

## References

Bartels, C. (2004): Acoustic correlates of 'second occurrence' focus: Towards an experimental investigation. In H. Kamp and B. Partee (eds.) *Context-dependence in the Analysis of Linguistic Meaning*. Amsterdam: Elsevier, 354–361.

Beaver, D., B. Clark, E. Flemming, T. F. Jaeger, and M. Wolters (2007) When semantics meets phonetics: Acoustical studies of second occurrence focus. *Language* 83. 245–276

Beckman, M. & J. Pierrehumbert (1986): Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255–309.

Berg, Rob van den, Carlos Gussenhoven & Toni Rietveld (1992) Downstep in Dutch: Implications for a model. In Gerhard Docherty & D. Robert Ladd (eds.) *Papers in Laboratory Phonology II: gesture, segment, prosody.* Cambridge: Cambridge University Press. 335–367.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.

Büring, Daniel (2001) Let's Phrase it! Focus, Word Order, and Prosodic Phrasing in German Double Object Constructions. In Gereon Müller and Wolfgang Sternefeld (eds.), Competition in Syntax. New York: Mouton de Gruyter, 69-105.

Büring, D. (2006): Focus projection and default prominence. In V. Molnár and S. Winkler (eds,) *The Architecture of Focus.* Berlin/New York: Mouton De Gruyter.

Büring, D. (this volume). Towards a Typology of focus realization.

Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24:239–298.

Féry, C. and L. Herbst (2004): German Sentence Accent Revisited. In Ishihara, S., Schmitz, M., Schwarz, A. (Eds.), *Working Papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS)* 1. Potsdam: Universitätsverlag Potsdam, 43–75.

Féry, C. and S. Ishihara (2009): The Phonology of second occurrence focus. Journal of Linguistics 45.

Féry, C. and F. Kügler (2008): Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36.4.

Féry, C. and V. Samek-Lodovici (2006): Focus projection and prosodic prominence in nested foci. *Language* 82.1. 131-150.

Féry, C. and H. Truckenbrodt (2005): Sisterhood and tonal scaling. *Studia Linguistica* 59. 223-243.

Fiedler, I., K. Hartmann, B. Reineke, A. Schwarz & M. Zimmermann (this volume) Subject focus in West African languages.

Fuchs (1976): 'Normaler' und 'kontrastiver' Akzent. *Lingua* 38, 293-312.

Gussenhoven, C. (1983): Focus, mode and the nucleus. *Journal of Linguistics* 19:377–417.

Gussenhoven, C. (1992): Sentence accents and argument structure. In I. Roca (ed.) *Thematic structure : its role in grammar*. Berlin & New York: Foris, 79–106.

Halle, M. & J. Vergnaud (1980): Three dimensional phonology. *Journal of Linguistic Research* 1: 83–105.

Hayes, B. (1995): *Metrical stress theory: Princieples and case theory*. Chicago: University of Chicago Press.

Hyman, L. & M. Polinsky (this volume) Focus in Aghem.

Ishihara, S. (2003): *Intonation and interface conditions.* Unpublished Ph.D. dissertation, MIT.

Ishihara, S. (2004): Prosody by phase: Evidence from the focus intonation–*wh*-scope correspondence in Japanese. In: S. Ishihara, M. Schmitz and A. Schwarz (eds.): *Interdisciplinary Studies of Information Structure: Working Papers of the SFB 632*, vol. 1. Potsdam: University of Potsdam, 77–119.

Ishihara, S. (2007): Major phrase, focus intonation and multiple spell-out (MaP, FI, MSO). *The Linguistic Review* 24: 137–167.

Ishihara, S. (in preparation): *Japanese downstep revisited*. Unpublished Ms.

Ishihara, S. and C. Féry (in prep) Syntactic and Prosodic Realization of Information Structure in German, Hungarian and Japanese.

Ito, Junko, and Armin Mester. 2006. Prosodic adjunction in Japanese compounds. In *Formal Approaches to Japanese Linguistics 4 MITWPL #55*. Cambridge, Mass.: MITWPL.

Ito, Junko, and Armin Mester. 2007. *Prosodic categories and recursion.* Paper presented at the Workshop on Prosody, Syntax, and Information Structure 3 (WPSI3).

Jackendoff, R. S. (1972) *Semantic interpretation in generative grammar.* Cambridge, Mass.: MIT Press.

Kanerva (1990): Focusing on Phonological Phrases in Chichewa. In: Inkelas, Sharon & Draga Zec (eds.) *Phonology-Syntax-Interface.* Chicago: University of Chicago Press. 145-161.

Katz, Jonah and Elisabeth Selkirk (2006). *Pitch and duration scaling for contrastive focus: a phrase stress analysis.* Unpublished manuscript, University of Massachusetts Amherst.

Kratzer, Angelika & Elizabeth O. Selkirk (2007) *Phase theory and prosodic spellout: The case of verbs*. *The Linguistic Review* 24: 93–135.

Kubozono, H. (1993): *The organization of Japanese prosody*. Tokyo: Kurosio.

Kubozono, H. (2006): Focus and intonation in Japanese: does focus trigger pitch reset? To appear in Ishihara (ed.) *Proceedings of the 2nd Workshop on Prosody, Syntax and Information Structure (WPSI2)*, *Interdisciplinary Studies on Information Structure* 8. Potsdam: University of Potsdam.

Kubozono, Haruo. (2007): Focus and intonation in Japanese: does focus trigger pitch reset? In S. Ishihara (ed.) *Proceedings of the 2nd Workshop on Prosody, Syntax and Information Structure (WPSI2), Interdisciplinary Studies on Information Structure 9*. Potsdam: University of Potsdam.

Ladd, R. (1990): Metrical representation of pitch register. In Kingston, J. & M. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press, 35–57.

Ladd, Robert D. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

McCarthy, J. and A. Prince (1993): Generalized Alignment. In Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology* 1993. Dordrecht: Kluwer. 79-153.

Nagahara, Hiroyuki (1994) *Phonological Phrasing in Japanese,* Doctoral dissertation, University of California, Los Angeles.

Nespor, M. & I. Vogel (1986): *Prosodic phonology*. Dordrecht: Foris.

Poser, W. (1984): *The phonetics and phonology of tone and intonation in Japanese.* Ph.D. dissertation, MIT.

Pierrehumbert, J. (1980): *The phonology and phonetics of English intonation.* PhD dissertation, MIT.

Pierrehumbert, J. and M. Beckmann (1988): *Japanese tone structure*. Cambridge: MIT Press.

Poser, William J. 1984. The phonetics and phonology of tone and intonation in Japanese. Doctoral Dissertation, Massachusetts Institute of Technology.

Rooth, M. (1985): *Association with focus.* Ph.D. dissertation, University of Massachusetts, Amherst.

Rooth, M. (1992): A theory of focus interpretation. *Natural Language Semantics* 1, 75–116.

Rooth, M. (2004): Comments on Krifka's paper. In: H. Kamp and B. Partee (eds.): *Context-dependence in the Analysis of Linguistic Meaning*. Amsterdam: Elsevier, 475–487.

Rooth, M. (this volume). Second occurrence focus and relativized F.

Schmerling, S. (1976): *Aspects of English sentence stress*. Austin: University of Texas Press.

Schwarzschild, Roger. 1999. Givenness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7:141–177.

Selkirk, E. (1984): *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge: MIT Press.

Selkirk, E. (1986): On derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.

Selkirk, Elisabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In *The handbook of phonological theory*, ed. Jerry A. Goldmsith, 550–569. Oxford, UK: Blackwell.

Selkirk, Elisabeth. 2002. Contrastive FOCUS vs. presentational focus : Prosodic evidence from right node raising in English. In Speech Prosody 2002. Aix-en-Provence, France.

Selkirk, Elisabeth. (2006). Bengali intonation revisited: An optimality theoretic analysis in which FOCUS stress prominence drives FOCUS phrasing. In Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation, C. Lee, and D. Büring (eds.). Dortrecht: Springer, 215–244.

Selkirk, E. (2008): Contrastive focus, givenness and unmarked status of "discourse new". In *The Notions of Information Structure,* C. Féry, G. Fanselow, and M. Krifka (eds.) *Acta Linguistica Hungarica* 125–145.

Selkirk, E. and K. Tateishi (1991): Syntax and downstep in Japanese. In C. Georgopoulos and R. Ishihara (eds.) *Interdisciplinary approaches to language: Essays in hornor of S.-Y. Kuroda*. Dordrecht: Kluwer Academic Publishers, 519–543.

Shinya, Takahito. 1999. Eigo to nihongo ni okeru fookasu ni yoru daunsteppu no sosi to tyooon-undoo no tyoogoo [The blocking of downstep by focus and articulatory overlap in English and Japanese]. In *Proceedings of Sophia Linguistics Society*, volume 14, 35–51.

von Stechow, A. and S. Uhmann (1986): On the focus-pitch accent relation. *Groninger Arbeiten zur Germanistischen Linguistik* 25, 228-263.

Sugahara, M. (2003): *Downtrends and post-FOCUS intonation in Japanese*. Ph.D. dissertation, University of Massachusetts, Amherst.

Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and prominence. Doctoral Dissertation, Massachusetts Institute of Technology.

Truckenbrodt, H. (2002): Upstep and embedded register levels. *Phonology* 19: 77–120.

Truckenbrodt, H. (2004): Final lowering in non-final position. *Journal of Phonetics* 32: 313–348.

Truckenbrodt, H. (2006) Phrasal stress. In The Encyclopedia of Languages and Linguistics, 2nd Edition, Keith Brown (ed.), Oxford: Elsevier, Vol. 9, 572-579.

Wagner, M. (2005): *Prosody and recursion*. Unpublished Ph.D. dissertation, MIT.