

Mayumi Hosono
University of Durham

On Phonological Feature Assignment*

1. Introduction. There are two major points of view on phonological features. One originates in a traditional assumption that phonological forms are registered in the lexicon. When lexical items are drawn from the lexicon and enter a syntactic derivation, phonological features are introduced with syntactic/semantic features, though the phonological features do not affect the syntactic operation. The phonological features are then stripped away from the syntactic object at a derivational point called Spell-Out (e.g. Chomsky 1995). The other is late insertion of phonological features. A syntactic derivation proceeds only with syntactic and semantic features. Only after the syntactic object is sent to Spell-Out, are phonological features introduced by morphological operations. This system is called Distributed Morphology proposed by Halle and Marantz (1993). Chomsky seems to suppose that results of syntactic derivations do not differ, whether the traditional view or Distributed Morphology is assumed (Chomsky 2000:119). There have been few discussions which deal with (dis)advantages of two approaches. I suspect, however, that the results will be different depending on which approach is taken, as I argue in the following sections. The aims of this paper are as follows. i) To show that (the derivational system based on) the traditional view does not work well, taking *uniformity* to be norm: assumption that the narrow syntactic and semantic components are uniform across languages (with surface appearance attributed to the phonological component), and assumption that a chain must be uniform (Chomsky 2000, 2001, 2004)¹. ii) To organize late insertion of phonological features within (feature system of) the current framework (Chomsky 2000~), proposing a new derivational model: it seems to have been unclear how syntactic/semantic features referred to in Distributed Morphology are dealt with in the current system (Chomsky 2000~), where do they originate from, are they drawn from the universal feature set or from a language-particular subset, and so forth. iii) To provide accounts for the issues which I claim appear to be problematic from the traditional view, based on the proposed model. In section 2 I discuss stripping-away of phonological features at Spell-Out, the traditional view assumed so far. I firstly discuss the lexical interface. I argue that assumption that idiosyncracies among languages lie in the lexicon of each language, which has been assumed from the early period of generative grammar (Chomsky 1981, 1995, Borer 1984), will not ensure uniformity of the narrow syntactic and semantic components. Next, I turn to the phonological interface. I argue the following points: i) timing of Spell-Out and a position from which phonological features are stripped away should be determined both by convergent narrow syntactic operations until Spell-Out and by the principles which lead a derivation to a convergent narrow syntactic component; therefore, presence/absence of (uninterpretable) phonological features will be irrelevant to whether narrow syntactic derivations converge or crash, unlike Chomsky (2000~); and ii) proposed mechanisms of

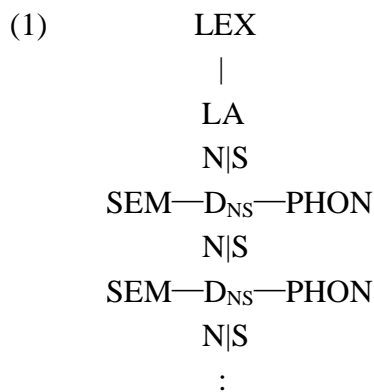
phonological feature assignment to a chain (Chomsky 2004, Nunes 1999, 2004) do not seem to work well, with the identical nature of occurrences taken into account. In section 3 I discuss late insertion of phonological features introducing Distributed Morphology proposed by Halle and Marantz (1993), which my proposal and analyses are based on. I point out that since source of syntactic/semantic features referred to in Distributed Morphology, that is where they originate from, is unclear, interaction between the other syntactic systems like narrow syntax and those features is also unclear. I assume that the features are derived in the universal set $\{F\}$. I propose a derivational model in which the features drawn from $\{F\}$ directly enter narrow syntax, with intermediate stages $[F] \rightarrow$ the lexicon \rightarrow lexical array assumed in Chomsky (2000~) eliminated: $\{F\} \rightarrow [F] \rightarrow$ ~~the lexicon~~ \rightarrow ~~lexical array~~ \rightarrow narrow syntax. I further propose to assume that the lexicon, a mechanism combining semantic/morphosyntactic features with phonological features, works after a narrow syntactic object is spelled out, namely, at the phonological component, which results in a model $\{F\} \rightarrow$ narrow syntax \rightarrow the phonological component (where the lexicon works). With the model, I turn again to the problems that are raised in the traditional view above. I emphasize that uniformity of the narrow syntactic and semantic components is strictly maintained, with surface differences which apparently belong to properties of individual lexical items as well as ones which are yielded from a different spelled out position in a chain all attributed to operations at the phonological component. I discuss phonological feature assignment to a chain in detail. Assuming uniformity of a chain and identical nature of occurrences, I claim that uninterpretable features should be deleted from all the occurrences in any chains before they are spelled out. I argue, based on late insertion of phonological features, that the point will shift from which position in a chain phonological features are stripped away from, to which position in the chain phonological features are inserted into. Based on the literatures which argue that features like [Agr] or [Foc(us)] can be introduced after Spell-Out (e.g. Halle and Marantz 1993, Erteschik-Shir 2001; cf. Rizzi 2004), I propose that one of the features enters a chain after Spell-Out and determines an actually pronounced position in the chain. In section 4 I briefly conclude this paper.

In the rest of this section I introduce theoretical backgrounds, mainly from Chomsky (2004). Except when I add words, I assume all below. A derivation constructs a pair $\langle \text{PHON}, \text{SEM} \rangle$.² *PHON* is accessed by the sensorimotor system *SM*; *SEM* by the conceptual-intentional system *C-I*. There is no interaction between *PHON* and *SEM* (Chomsky 2004:110). The *Interface Condition* is imposed on a derivation: the information in the expression generated in narrow syntax *NS* must be legible to other cognitive systems that enter thought and action. The derivation *converges* if *PHON* and *SEM* each satisfy the Interface Condition; otherwise, the derivation *crashes* (Chomsky 2004:106). The economy principles since Chomsky (1995) are also imposed on a derivation: any superfluous elements in representations and any superfluous steps in derivations are eliminated (Chomsky 2000:99).

Initial state S_0 determines the set of linguistic features $\{F\}$; from the set a subset $[F]$ is drawn, which is assembled to make the *lexicon* *LEX* in a language; from *LEX* a lexical array *LA* is accessed and selected in each derivation *D* (if accessed more than once, a *numeration*); *LA* enters an *NS* derivation (Chomsky 2004:107). *LA* is assumed to be extended to

numeration: after LA is selected from LEX, a subarray LA_i is selected from LA at each phase to reduce a computational burden (Chomsky 2000:106).³ I will propose a syntactic model in which intermediate stages from [F] to $LA_{(i)}$ are eliminated with features in {F} directly entering an NS operation.

Each language has the following three components: NS maps LA to a derivation D_{NS} ; the phonological component Φ maps a D_{NS} to PHON; and the semantic component Σ maps D_{NS} to SEM. NS and Σ are assumed to be *uniform* for all languages, while Φ is highly variable and LEX is the locus of parametric difference among languages (Chomsky 2004:107). I will argue that it is only PHON that is differed among languages, assuming that a mechanism which associates semantic/morphosyntactic features with phonological ones (i.e. LEX) works at PHON. Mappings will satisfy the Inclusiveness Condition, introducing no new elements but only rearranging them in a domain. Φ and Σ apply to units constructed by NS and ‘proceed cyclically in parallel’ (Chomsky 2004:107). Mappings to Φ is called *Spell-Out* S-O, which ‘removes from NS all features that do not reach SEM. ... [W]e refer to all these as “phonological”’ (Chomsky 2004:125,ft.14). S-O applies cyclically: an NS product is sent to PHON and SEM phase-by-phase, that is in a local way (Chomsky 2004:107). The current model is roughly illustrated as follows:



A single output model until Chomsky (1995), which has only one S-O as a point where an NS component is sent to SEM and PHON respectively, has been abandoned since Chomsky (2000). I will advocate that such a single level model is preferable, though I do not assume the independent SEM component, based on late insertion of phonological features (Halle and Marantz 1993).

NS proceeds with *Merge*: it takes two elements α and β and creates a new unit; it applies iteratively. $\{\alpha, \beta\}$ is a projection, identified either by α or by β (its label, which is always that of a head). The number of Specs are not limited, since the limitations on Merge follow from selectional and other conditions that are independent. Merge of α to β requires β to search the closest α under *c-command*, satisfying locality condition. Merge satisfies the *Extension Condition* (Chomsky 2004:108-109).

When α and β are separate objects, Merge is *external*; when one is part of the other, the operation is *internal*, yielding displacement. Internal Merge leaves a *copy*. The copy is in effect an *occurrence*, an entity identical with the other one. Later, I will discuss definition of

occurrences, and so forth, more in detail. Occurrences compose a *chain* in the following way: α is drawn from LEX as part of LA; α is further drawn and copied from LA as part of LA; when α ‘moves,’ α is once again copied from LA (i.e. extending LA tonumeration). Then, two α s forming a chain $\langle \alpha, \alpha \rangle$ are two occurrences of the same α (Chomsky 2000:114-115). Proposing that {F} directly enters NS, I will assume that α is a feature (complex) copied directly from {F} and that a chain α s form is composed by syntactic(/semantic) features. The copy is defined as follows: K is a copy of L if K and L are identical except that K lacks the phonological features of L. Application of internal Merge before S-O yields overt movement, with β in the pair $\langle \alpha, \beta \rangle$ losing the phonological features. Its application after S-O, on the other hand, yields covert movement, with α in $\langle \alpha, \beta \rangle$ losing the phonological features (Chomsky 2004:110-111). I will claim that internal Merge after S-O is impossible in any way. I will propose that a pronounced position is determined by a feature that can be introduced into a chain after S-O, based on Halle and Marantz (1993) and Erteschik-Shir (2001).

It is assumed that C, T,⁴ and v⁵ are the core functional categories, possessing uninterpretable ϕ -features. v and T are *probes*⁶ for the Case-agreement system. CP and v(*)P, but not TP, are propositional in that CP is a full clause containing tense and force, and v(*)P is a projection in which all θ -roles are assigned. CP and v(*)P are called *phases* (Chomsky 2000:102). Transitive v*P and CP are strong phases; intransitive vP is a weak phase. A typical phase is as follows: PH = [α [H β]] (α and H(ead) are edges of PH.) (Chomsky 2004:108). The *Phase Impenetrability Condition* states that ‘the domain of H is not accessible to operations, but only the *edge* of HP’ (Chomsky 2004:108). All operations simultaneously apply at the phase level; Spell-Out too applies at each phase (Chomsky 2004:123).

An extra Spec, a non- θ position, of T is allowed by the *Extended Projection Principle* EPP, which ‘might be universal’ (Chomsky 2000:109); pure Merge to non- θ position is restricted to [Spec,TP] for the θ -theoretical reason. On the other hand, extra Specs of C and v(*) are allowed by uninterpretable EPP-features when available, which optionality characterizes phases.⁷

NS has been assumed to have uninterpretable features as mechanisms that force displacement (Chomsky 2000~). Uninterpretable features – the EPP, structural *Cases* for nouns, ϕ -features of T for subject-agreement and those of v for object-agreement, for instance – must be eliminated before an NS derivation is sent to Σ . Uninterpretable features come into LEX without *values*, distinguished from interpretable features (Chomsky 2004:116). Proposing that {F} directly enters NS, I assume that each feature stands in a ‘primitive’ form. Taking *what* as example, I suppose as follows: i) a semantic feature [what] and an uninterpretable *wh*-feature [u-wh] stand by themselves in {F}; ii) [what] and [u-wh] are combined to be a unit [what]+[u-wh], which enters the following NS derivation. Uninterpretable features of α are eliminated in an appropriate relation to interpretable features of β that is complete with a full set of features. The procedure deleting uninterpretable features are made roughly as follows: a head, available without search, has uninterpretable features, a ϕ -set. The ϕ -set as a probe seeks the closest matching features (a *goal*), making it active; the uninterpretable ϕ -set of the probe matches the interpretable counterpart of the goal (*Match*). Feature matching is non-distinctness rather than identity. The matching operation

must be performed as quickly as possible, and prohibits a partial elimination of features: ‘Maximize matching effects’ (Chomsky 2001:15). The uninterpretable features are valued by the matching features in an operation called *Agree* and eliminated until S-O. If the ϕ -set of the probe also has the EPP, the goal, which has an uninterpretable Case and can still be active, selects a phrase, which moves to delete the EPP (*Move*) (Chomsky 2004:113-114). The head must have a complete set of ϕ -features (i.e. it must be *ϕ -complete*) to delete uninterpretable features in the *Agree* operation (Chomsky 2001:6). *Agree/Match* being assumed to apply freely, the probe-goal relation must be evaluated at the strong phase level. In the following configuration $\alpha > \beta > \gamma$ ($>$ indicates a c-command relation; both β and γ can match the probe α) if uninterpretable features of β is deleted, β is rendered inactive and unable to move to delete the EPP (‘frozen in place’ (Chomsky 2000:123)): the effects of matching between α and γ are blocked (*defective intervention constraint*). When there are phonological features at the outer edge of $v(*)P$ (i.e. the phonological edge) too, *Match* between the probe in the higher phase and the goal in the lower phase is prevented under the *Minimal Link Condition*.

2. Stripping-away of phonological features at S-O. It has long been assumed that language variation lies in LEX (and in PHON) since the early period of generative grammar (Chomsky 1981). LEX specifies phonological, syntactic, and semantic properties of each lexical item (Chomsky 1981, 1995). They are introduced into an NS derivation, though only syntactic features affect an NS operation. Phonological features are *stripped away* from an NS component at S-O (Chomsky 1995:229). A proposal on disposal of phonological features has been made: NS consists of syntactic/semantic features; phonological features are assigned only after S-O (*Distributed Morphology*, Halle and Marantz 1993). Chomsky states that though an output of an NS component will not differ between both approaches, late insertion of phonological features in DM requires a redundant stipulation that a “placeholder” F’ must be replaced with an F identical with F’ (Chomsky 2001:11). In this section I inquire into the traditional model which is still assumed in the current system (Chomsky 2000~), in which phonological features are specified in LEX and stripped away from NS components; I also discuss results which are derived from the model. I claim that the model is insufficient to account for several points which I illustrate below.

2.1. Interface with LEX. The approach which assumes phonological features to be stripped away at S-O is illustrated in the model above (1). A derivation starts from LA drawn from LEX; LEX feeds NS. This idea has never been changed since the beginning of generative grammar. LEX is a locus of parametric difference, while Σ SEM and NS are uniform for all languages (Chomsky 2004:107).⁸ LEX has been assumed to be a list of “exceptions” (Chomsky 1995:235). Optimal codings of idiosyncratic properties of each lexical item in a particular language are given in a unified entry. Therefore, each lexical entry contains information of formal features, information as instructions for PHON (i.e. a phonological matrix), and information for interpretations at SEM (i.e. semantic properties) (Chomsky 1995:238). LEX includes substantive categories like N, A, V, and P as well as functional categories like C, T, Agr, and D (Chomsky 1995:240).

The idea that parametric difference among languages lies in LEX originates in Borer (1984). She argues that parametric difference depends on i) whether a particular inflectional rule is available to a language and ii) at which level application of the inflectional rule is restricted (Borer 1984:27). She raises the following examples:

- (2) a. *hkit ma9-o la Karim*
 talked-I with-him to Karim
 ‘I talked with Karim.’
 (Lebanese Arabic, Borer 1984:27,(34c), from Aoun 1982)
- b. **dibarti ’im-a (le/šel) Anna*
 talked-I with-her to/of Anna
 ‘I talked with Anna.’
 (Modern Hebrew, Borer 1984:27,(35c))

Lebanese Arabic allows clitic doubling, which is illustrated by possible cooccurrence of *o* and *la Karim* in (2a); Modern Hebrew, on the other hand, does not allow clitic doubling (2b). Borer accounts for the difference as follows: Case property of prepositions *ma9* in Lebanese Arabic and *’im* in Modern Hebrew is absorbed by clitics *o* and *a* respectively. Lebanese Arabic, however, has a ‘saving device’ to assign Case to a complement NP: insertion of the preposition *la*. Modern Hebrew does not have such a device; therefore, ungrammaticality of (2b) arises (Borer 1984:28). She argues that the following inflectional rule is available to Lebanese Arabic, but not to Modern Hebrew:

- (3) $\emptyset \rightarrow la / [_{PP} \dots NP]$
 (Borer 1984:28,(37))

Based on the argument above, Borer claims that parametric variation affects only inflectional system, thus individual lexical items associated with functional categories (Borer 1984:29).

Assuming uniformity of NS and SEM for any language (Chomsky 2001, 2004), however, doubt should be thrown on whether uniformity is ensured on the assumption that LEX can differ among languages. Sigurðsson (2003) argues that contradiction will arise between uniformity and feature selection from the universal set {F}. The Uniformity Principle says, ‘In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances’ (Chomsky 2001:2). On the assumption that a language selects particular sets of features from {F} as its property [F] (Chomsky 2000~), all languages should be able to access {F}. Assume i) that {F} contains {F₁, F₂, F₃, F₄, F₅}; and ii) that L₁ selects {F₁, F₃, F₄} for its [F] while L₂ selects {F₂, F₃, F₅} for its [F]. This will produce contradiction to the Uniformity Principle: L₁ and L₂ would not access {F₂, F₅} and {F₁, F₄} respectively, though each language would access all features in {F} (Sigurðsson 2003:325-326).

I make case studies of the issue with specific examples. Let us take functional categories T

and Agr, and account for how derivations would differ given different LAs.⁹ See below:

- (4) a. Il telefonerà.
he telephone-FUT-3sg
'He will telephone.'
- b. kare-ga denwasuru-darou.
he-Nom telephone-FUT
'He will telephone.'

Meaning (i.e. a SEM output) of (4a) (Italian) and (4b) (Japanese) is not different, which is shown in the translations. Thus, it would be expected that the SEM outputs are produced with the same NS derivations in both languages. However, since Italian has rich agreement system, T as well as Agr should be contained in LA, resulting in LA = {Il, telefonerà, T, Agr} (4a). Japanese has no agreement system; therefore, only T would be included in LA, resulting in LA = {kare-ga, denwasuru-darou, T} (4b). This leads to a situation in which LAs drawn from LEX (i.e. NS inputs) differ between Italian and Japanese. Assume that NS has produced [T [il telefonerà]] for Italian (4a) and [T [kare-ga denwasuru-darou]] for Japanese (4b) after a series of Merge operations.¹⁰ They both are derived from the same, 'uniform' operations until this stage. Agr is still left for Italian; one more operation which merges Agr to [T [il telefonerà]] must occur in the Italian NS. If Agr were assumed to be merged before T, order of Merge would further be different between Italian and Japanese.¹¹ NS outputs result in [Agr [T [il telefonerà]]] for Italian and [T [kare-ga denwasuru-darou]] for Japanese. In this way, different LAs (i.e. different NS inputs) will be expected to produce different NS outputs. Consequently, different LAs will not ensure uniformity of NS (, though, fortunately, SEM outputs do not differ in this case due to Agr's semantic emptiness (Chomsky 1995)).

It is also important to discuss a case in which parametric difference is assumed to lie in selection property of functional categories, though LAs are not different. According to Ouhalla (1991), selection property is crucial to different order possibility of merging functional heads. Based on Borer's (1984) line, he states that 'a given functional category may select a specific category in one language and a different one in another, thus giving rise to a difference in the arrangements of these categories in the structure' (Ouhalla 1991:8). He proposes Agr/T parameter: T c-selects Agr in VSO languages, while Agr c-selects T in SVO languages (Ouhalla 1991:113). See below (I slightly modified.):

- (5) a. sa-ya-shtarii Zayd-un dar-an.
FUT(T)-3sg.MASC(Agr)-buy Zayd-Nom house-Acc
'Zayd will buy a house.'
- b. legge-va-no.
read-(T)-3pl.(Agr)
'They read.'

Assuming that basic order in Arabic is VSO, Agr is inside T (5a). Assuming that an Italian verb moves to Agr, T is inside Agr (5b) (Ouhalla 1991:113-114). Arabic and Italian both contain T and Agr: LAs as NS inputs are not different. Assuming that selection property of those functional categories differs among Arabic and Italian, NS operations will proceed in a different way: Agr is always firstly merged, and T secondly, in VSO languages like Arabic; on the other hand, T is firstly merged, and Agr secondly, in SVO languages like Italian. Here again, uniformity of NS components seems to be difficult to be maintained, assuming idiosyncrasies to lie in selection property of functional heads: even if LAs do not differ, resulting NS outputs would be differentiated.

Consequently, both the assumption that idiosyncratic properties lie in LEX and the assumption that parametric difference among languages is attributed to selection property of functional heads will not ensure uniformity of NS (and maybe SEM too) for all languages. Namely, if LAs as NS inputs are different, NS operations as well as NS outputs will differ; if NS outputs, in other words SEM inputs, differ, SEM outputs may also be differentiated.

2.2. Interface with PHON. Let us turn to the interface with PHON. In addition to LEX, Φ -PHON is also assumed to be a locus of high idiosyncrasies among languages, unlike NS and Σ -SEM that are uniform (Chomsky 2004:107). It has been a standard assumption that a lexical entry of each lexical item contains phonological features in addition to formal and semantic features (Chomsky 1995:238). Phonological features are uninterpretable (Chomsky 2001:4); they must be stripped away from an NS object at S-O and sent to PHON in order for a derivation not to crash at SEM (Chomsky 2000:118); not phonological features themselves, but only their presence/absence, can affect NS derivations (Chomsky 2001:10, 2004:ft.64).¹² I throw doubt on some points of the assumptions in turn.

2.2.1. Presence/absence of phonological features. Firstly, I would like to discuss whether presence/absence of phonological features affect NS derivations. See below:

- (6) a. Who said what?
 b. *What did who say t_{what} ?

(6a-b) illustrate the superiority effects (Chomsky 1995, Richards 1997, Pesetsky 2000). Based on Chomsky (2000), derivations of (6a-b) proceed in the following way. C that possesses uninterpretable features [u-Q] merges to TP, resulting in C [_{TP} who T [_{v*P} t_{who} say what]]. C with [u-Q], a probe, seeks a goal that has its interpretable counterpart [Q] to delete [u-Q]. A candidate is either *who* or *what*. *Who* is chosen, because it is the closest category with [Q] from the probe C. Match takes place between C and *who*; C's [u-Q] is deleted by *who*'s interpretable [Q]. C has the EPP too; it must be deleted by a category which is activated with some uninterpretable feature the category itself has. In this case a candidate is either *who* or *what*, both of which have an uninterpretable feature [u-wh]. *What* cannot be chosen, crossing *who*, which causes the defective intervention effect as in (6b). Consequently, only *who* can be

selected as a category that deletes C's EPP. The phonological features of *what* (and *said*) are stripped away at the original position, while those of *who* are stripped away at [Spec,CP].¹³

Crucial is that only *who*, but not *what*, can be chosen as the candidate that deletes C's EPP: it is (6a), not (6b), that can be constructed in the NS operations before S-O and sent to PHON. It appears that a position from which phonological features are stripped away and timing of spelling out the position are determined by convergent NS operations before S-O like Match and Move as well as by the principles which lead to the convergent NS component: to avoid the defective intervention effect, *who* must be selected and move to [Spec,CP], which constructs the NS component [who [said what]]; the phonological features of *what/said* are stripped away at the original positions, while those of *who* at [Spec,CP]. It might be argued that (6b) is ungrammatical because the phonological features of *who* are present in [Spec,TP]. This argument presupposes that *who* is spelled out in [Spec,TP]. As long as *who* must move to [Spec,CP] to avoid the defective intervention effect, S-O of *who* in [Spec,TP] will be prevented in a principled way. Therefore, it does not seem to be a case that presence/absence of phonological features affects NS derivations; rather, their presence/absence in a certain position, in other words timing of S-O and a position from which phonological features are stripped away, should all be determined by convergent NS operations and the principles that lead the NS component to converge.

2.2.2. Phonological feature assignment to a chain. Next, I would like to make mention of phonological feature assignment to a chain. See below:

- (7) a. what did you eat _{t_{what}}?
- b. kimi-wa nani-o tabemashi-ta-ka?
 you-Top what-Acc eat -past-Q
 'what did you eat?'

What of English is situated in a sentence initial position (7a), while *nani-o* of Japanese remains in an original position (7b). *Wh* moves to delete its own [u-wh] (Chomsky 2004:115). Therefore, English and Japanese each has a chain which consists of occurrences of *whs*, <what,what> and <nani-o,nani-o> respectively.¹⁴ The only difference between English and Japanese is whether *wh* is pronounced in a higher position (English) or in a lower position (Japanese). Assuming that a chain is composed of copies/occurrences identical with each other (Chomsky 2000), a question will arise: how is a pronounced position in a chain determined?

Let us see first Chomsky's (2004) proposals on movement and chain formation. Movement leaves copies, that is occurrences identical with each other and forming a chain (Chomsky 2000~). Assuming that *wh* moves to delete its own [u-wh] after its Case is deleted in Agree with v* (Chomsky 2000~), chains of *what* and *nani-o* in (7) are respectively as follows:

- (8) a. [_{CP} what₁ did [_{TP} you [_{v*P} what₂ [_{v*P} eat what₃]]]]¹⁵

- b. [CP nani-o₁ [TP kimi-wa [v*P nani-o₂ [v*P nani-o₃ tabemashi-ta-ka]]]]¹⁶

On the assumption that application of internal Merge before S-O yields overt movement while its application after S-O yields covert movement (Chomsky 2004:110-111), *what* is spelled out after it has internal-merged to [Spec,CP] (i.e. at the position of *what*₁), resulting in overt movement and two chains, <what₂,what₃> and <what₁,what₂> (7a); *nani-o* internal-merged to [Spec,CP] after it is spelled out in situ (i.e. at the position of *nani-o*₃), resulting in covert movement and two chains, <nani-o₂,nani-o₃> and <nani-o₁,nani-o₂> (7b).

Supposing that *wh* moves even after its phonological features are stripped away at S-O, the following example would be predicted to be grammatical as a normal *wh*-interrogative:

- (9) You ate what?

(9) is interpreted only as an echo question. Assuming internal Merge after S-O, nothing would prevent *what* from moving covertly to [Spec,CP] to delete its own [u-wh] after its Case is deleted in an Agree operation with v*, contrary to the fact. It would be necessary to say that internal Merge must apply before S-O in English, while it must apply after S-O in Japanese.¹⁷

Further, I would like to throw doubt on feasibility of internal Merge after S-O. I repeat a Japanese *wh*-chain below:

- (10) [CP nani-o₁ [TP kimi-wa [v*P nani-o₂ [v*P nani-o₃ tabemashi-ta-ka]]]]

After Agree takes place between v* and *nani-o*, Case of the latter is deleted. It might be said that it is only the chain <nani-o₂,nani-o₃> that is related to [u-Case]; [u-wh] is solely involved in the chain <nani-o₁,nani-o₂>. Thus, it could be argued that after Case deletion *nani-o* can be spelled out in situ, after which *nani-o* covertly moves up to [Spec,CP] to delete [u-wh]. However, it seems to be difficult to assume that neither *nani-o*₂ nor *nani-o*₃ is involved in [u-wh], though. It will not be plausible to suppose that [u-wh] is not attached to *wh* at the numeration, but it enters in the course of derivation after *wh* is spelled out: [u-wh] is arguably *wh*'s inherent feature that characterizes *wh* as an operator. Namely, two occurrences of *nani-o* in <nani-o₂,nani-o₃> will have [u-wh] both before and after [u-Case] is deleted. Supposing that an NS component is spelled out only after uninterpretable features are eliminated, the chain <nani-o₂,nani-o₃> should not be spelled out: unless [u-wh]s are deleted from the *wh*-occurrences, the chain would not be a legitimate syntactic object, which would cause a derivation to crash.

To assume that an operation, in this case *wh*-movement, applies only at a phonological edge of a phase, with S-O applying to the complement of the head of the phase (Chomsky 2004:12) will not save a derivation, either. In (10) *nani-o*₂ in [Spec,v*P] and *nani-o*₃ in the original position are occurrences that form a chain. Assuming *nani-o* moves to [Spec,v*P] with [u-wh] before S-O, it would be argued that only *nani-o*₂, but not *nani-o*₃, has [u-wh]; therefore, *nani-o*₃ without [u-wh] could be spelled out in situ. Assuming uniformity of a chain, however, the *wh*-chain consisting of *nani-o*₂ and *nani-o*₃ is not uniform: *nani-o*₂ has [u-wh], though

nani-o₃ may not have any. Supposing that uninterpretable features are deleted from a set of occurrences, namely from the whole chain (Chomsky 2000:116), *nani-o₃* cannot be assumed to have no [u-wh]: as long as *nani-o₃* is contained in the *wh*-chain that includes the occurrence with [u-wh], the former will surely share [u-wh]. Since the chain is not a legitimate syntactic object, it cannot be spelled out: S-O of the chain would lead a derivation to crash.

It could alternatively be assumed that a final chain only has to satisfy all the conditions like uniformity though a chain in an intermediate derivational stage does not have to. This will not save a derivation, either. See below:

- (11) [CP *nani-o₁* ... [_{v*P} *nani-o₂* [_{v*P} ... *nani-o₃* ...]]]
 what what what

Assume that [u-wh] is deleted in the final position [Spec,CP]. Since [u-wh], once deleted, is eliminated from all the *wh*-occurrences, two chains, <*nani-o₂*,*nani-o₃*> and <*nani-o₁*,*nani-o₂*>, will both be legitimate after [u-wh] deletion in that the chains are uniform. Therefore, S-O of the chains would have no problem. However, it is only after [u-wh] is deleted in the position of *nani-o₁* that the chains would not have contained [u-wh]; it is only at this point of derivation that S-O of the chain would be allowed without having any [u-wh]. Thus, even if it were already known that [u-wh] is deleted in the topmost position, S-O of *nani-o₃* would not be allowed before the derivation reaches the position: [u-wh] could not be deleted from the *wh*-occurrences at any stage of the derivation before *nani-o* reaches [Spec,CP]. Consequently, S-O in the course of a derivation, in other words internal Merge after S-O, appears to be impossible in any way, on the assumption that a chain must be uniform.

Let us turn to Nunes (1999, 2004). Unlike the approach proposed by Chomsky above, in which a pronounced position in a chain must be stipulated in some way, he attempts to determine the pronounced position in a principled way. See below:

- (12) John was kissed (*John).

All copies, either a head (an intermediate copy, if any) or a tail, should be subject to the same principle under the assumption that copies are nondistinct (Nunes 2004:16): in (12) both *John* in a higher position and *John* in a lower position should not be prevented from being pronounced. If one of the copies is not deleted in a chain, a syntactic object cannot satisfy the *Linear Correspondence Axiom* LCA, which employs a notion of asymmetric c-command to determine word order (Kayne 1994). That is, on the assumption that both *Johns* above are nondistinct, the higher *John* would asymmetrically c-command, and the lower *John* would asymmetrically be c-commanded by, the Aux *was* (Nunes 2004:24). Therefore, deletion of chain link(s) is required for *linearization* in accordance with the LCA (Nunes 2004:25). With the argument that formal features are relevant to computations at PHON, deletion of uninterpretable formal features renders them invisible not only at SEM, but also at PHON (Nunes 2004:32). In the example above (12) a relevant uninterpretable feature is *John's* Case. The structure is represented as follows:

(13) [John-CASE was kissed John-CASE]

Case is deleted in [Spec,TP] in a Case-deleting relation with the matrix T, as small capitals illustrate. With *Chain Reduction*, which says, ‘delete the minimal number of constituents of a nontrivial chain CH that suffices for CH to be mapped into a linear order in accordance with the LCA’ (Nunes 2004:27), the most economical way is delete one copy. Possible patterns are given below:

(14) a. [John-CASE was kissed ~~John-CASE~~]

b. [~~John-CASE~~ was kissed John-CASE]

Since Case is already deleted in [Spec,TP], (14a) requires no further operation. (14b), on the other hand, requires one more operation of deleting Case from the lower position to lead a derivation to converge, as follows:

(15) [~~John-CASE~~ was kissed John-CASE]

Compared with (15), (14a) is derived in a more economical way than the former. Consequently, (14a) is determined as an output at PHON: a phonetically realized output is determined in the way that formal features are deleted in the most economical way (Nunes 2004:32-33).

I would like to raise several questions in Nunes’ system. First, his notion of economy sounds to be strange. Nunes states as follows:

Exploring the null hypothesis regarding the copy theory of movement, the above proposal thus takes the position that both heads of chains and traces should in principle be subject to phonetic realization. According to the logic of the proposal, there is nothing intrinsic to lower copies that prevents them from being pronounced. If Chain Reduction proceeds in such a way that only a trace survives, the derivation may eventually converge at PF. The fact that in most cases such a derivation yields unacceptable sentences is taken to follow from *economy* considerations, rather than convergence at PF. Since the highest chain link is engaged in more checking relations, it will require fewer application of F[ormal]F[eature]-Elimination than lower chain links, thereby being the optimal candidate to survive Chain Reduction and be phonetically realized, all things being equal (Nunes 2004:33).

As is obviously shown above, he presupposes that *S-O in a higher position is unmarked* comparing with S-O in a lower position, though nothing ‘prevents [a lower copy] from being pronounced.’ According to him, this ‘follow[s] from economy considerations.’ He claims that multiple *wh*-interrogative is accounted for with this system. See German examples below:

- (16) a. *Wen denkst Du wen sie meint wen Harald t_{wen} liebt?*
who think you who she believes who Harald t_{wen} loves
 ‘Who do you think that she believes that Harald loves?’
- b. **Wen glaubt Hans wen Jacob wen gesehen hat?*
whom thinks Hans whom Jacob whom seen has
 ‘Who does Hans think Jacob saw?’
 (Nunes 2004:39,(75-76), originally from Fanselow and Mahajan 1995; I slightly modified.)

In (16a) only the head and the intermediate *wh*-copies are phonetically realized; in (16b), on the other hand, the tail is realized too. The chains are represented as follows:

- (17) a. [wen ... [wen ... [wen ... ~~wen~~]]]
- b. *[wen ... [wen ... wen]]

With Chain Reduction applied, the difference in (un)grammaticality is accounted for roughly as follows. First, when a chain is not linearized in accordance with the LCA, member(s) of the chain must be subject to reduction; therefore, not all the copies can be phonetically realized, as (17b) illustrates. Second, when Chain Reduction is necessary for linearization in accordance with the LCA, as small number of members as possible should be deleted; therefore, deletion of only one *wh*-copy is justified, as illustrated in (17a) (Nunes 2004:41-42).

Compare the German case with an English counterpart expressed in the translation. A chain representation of the English counterpart of (16a) is as follows:

- (18) a. Who do you think that she believes that Harald loves?
- b. [who ... [~~who~~ ... [~~who~~ ... ~~who~~]]]

English does not pronounce the intermediate *wh*-copies as illustrated in (18). Following Nunes, a derivation in English would be less economical than that in German, since more chain links are reduced in English than in German. Consider economy of articulation, however, which is omnipresent in human languages: language prefers less and/or shorter expressions to more and/or longer ones, as illustrated in ellipsis or omission. Thus, it will not be plausible that a language/sentence construction that pronounces more numbers of a chain is more economical than a language/sentence construction that has less phonetic realization of chain members; rather, the less phonetic realization, the more economical phonetic computation of human language will be.

In addition, how about languages like Chinese and Japanese in which *S-O in a lower position is unmarked*? It is well-known that a *wh*-copy is always realized in situ in the

languages (e.g. Huang 1982, Watanabe 1992). I repeat a Japanese example below:

- (19) kimi-wa nani-o tabemashi-ta-ka?
 you-Top what-Acc eat -past-Q
 ‘what did you eat?’

A chain consisting of occurrences of *nani-os* is represented as follows:

- (20) [~~nani-o~~ ... [nani-o ...]]

Let us turn to Nunes’ statement again: ‘if Chain Reduction proceeds in such a way that only a trace survives, the derivation may eventually converge at PF. The fact that in most cases such a derivation yields unacceptable sentences is taken to follow from *economy* considerations ...’ From the data of the languages like Japanese, it is definitely said that it is *not* ‘the fact that in most cases [a] derivation [in which ‘only a trace survives’] yields unacceptable sentences,’ contrary to his claim. Regarding this point, Nunes’ notion of economy sounds to be strange to a native speaker of the languages which have S-O in-situ as an unmarked option.

He seems to attempt to account for in-situ S-O cases (i.e. covert movement) in terms of *sideward movement* of formal features: “‘covert feature movement’ can be reanalyzed as overt sideward movement of F[ormal]F[eature]s’ (Nunes 2004:153). Raising covert head movement as in English, he argues as follows. After VP is generated, V’s formal features sideward-move and adjoin to T, resulting in two syntactic objects [_{VP} ...Vⁱ...] and [_T FF(Vⁱ)+T]; they merge, resulting in [_{TP} [_T FF(Vⁱ)+T][_{VP} ...Vⁱ...]]; a resulting chain of V’s formal features would be <FF(Vⁱ),FF(Vⁱ)>. Since the chain consists of nondistinct copies, they cannot form a chain; therefore, they are not subject to Chain Reduction (Nunes 2004:153-154).¹⁸

It seems to be doubtful whether the account in terms of formal feature movement applies to *wh*-movement, since *wh*-movement includes [u-wh] deletion. Based on Nunes’ argument, *wh*’s FF would be attached to an interrogative head C, which would make the Japanese *wh*-chain (21) like [CP FF(whⁱ)+C ...[whⁱ ...]]. FF(whⁱ), however, must be different from the original whⁱ, since [u-wh] is deleted in the higher scope position; therefore, a resulting chain will be [CP FF(WHⁱ)+C ...[(WHⁱ) ...]] (i.e. <FF(WHⁱ),FF(WHⁱ)>). As long as the chain is formed by distinct copies, the chain must be subject to Chain Reduction; it is not clear how an account would continue under his system.

Alternatively, let us tentatively account for *wh*-in-situ in terms of category movement, as Nunes does for overt *wh*-movement. A derivation of the Japanese *wh*-interrogative (19) will proceed as follows:¹⁹

- (21) a. [nani-o-WH kimi-wa nani-o-WH tabemashi-ta-ka]
 what you what eat -past-Q
 b. [~~nani-o-WH~~ kimi-wa nani-o-WH tabemashi-ta-ka]

- c. [~~nani-o-WH~~ kimi-wa nani-o-WH tabemashi-ta-ka]

Assume that [u-wh] is deleted in the topmost position based on Chomsky (2000) (21a). Assume further that Chain Reduction applies to a higher position in Japanese (21b). The derivation of the syntactic object would crash, though: [u-wh] remains to be deleted in the tail. Thus, one more operation deleting the [u-wh] is required, which results in (21c). Following Nunes, (21c) is an NS output less economical than one resulting from a derivation of the English counterpart in which Chain Reduction applies only to a tail position, though SEM outputs are not different between the languages. It sounds to be strange to say that a derivation in the languages like Japanese is always less economical than others, considering the fact that an in-situ strategy is an unmarked option in the languages.

It could be assumed i) that [u-wh] is deleted in a head position in languages with overt movement, but in a tail position in those with *wh*-in-situ, and ii) that Chain Reduction applies to a tail position in the former, but to a head position in the latter:

- (22) a. What did you eat t_{what} ?

- b. kimi-wa nani-o tabemashi-ta-ka?
 you-Top what-Acc eat -past-Q

- (23) a. [what-WH ... ~~what-WH~~]

- b. [~~nani-o-WH~~ ... nani-o-WH ...]

Since the number of required operations would not be different, a derivation would result from the most economical way in both languages. There is no reason, however, to assume that [u-wh] is deleted in different positions among languages; rather, it will only be a stipulation, an undesirable situation.

Second, in relation to the discussion just above, it is doubtful that uniformity of NS and/or SEM (Chomsky 2004) is maintained based on Nunes' system. I repeat the relevant examples below:

- (24) a. what did you eat t_{what} ?

- b. [what-WH did you eat ~~what-WH~~]

- (25) a. kimi-wa nani-o tabemashi-ta-ka?
 you-Top what-Acc eat -past-Q
 'what did you eat?'

- b. [~~nani-o-WH~~ kimi-wa nani-o-WH tabemashi-ta-ka]

(26) a. You saw what?

b. [~~what-wh~~ you saw what-~~WH~~]

(24) looks like the most economical derivation: [u-wh] is deleted in [Spec,CP]; Chain Reduction takes place only in the lower position. In (25) the higher position is subject to Chain Reduction. To avoid crashing at PHON, however, one more operation deleting [u-wh] in the lower position is required. In (26) too Chain Reduction occurs in the higher position; further operation is still required to delete the lower [u-wh]. First, comparing (24) with (25), the point is that (24) is semantically equivalent to (25), though the required NS operation of the former is different from that of the latter: (25) requires one more operation than (26). This is a case in which NS is not uniform, though SEM is (or happens to be) non-distinct. It could be assumed that [u-wh] is deleted in a head position in languages with overt movement, but in a tail position in those with *wh*-in-situ; this, however, would lead to a stipulation, as stated previously. Next, comparing (25) with (26), the point is that (25) is not logically equivalent to (26), though the required number of NS operations are not different between them. This is a case in which though NS is uniform, SEM differs. Even if formal feature movement is assumed for in-situ S-O cases (25-26), the results are not different. Consequently, nontrivial numbers of derivation which do not maintain uniformity of NS and SEM seem to be produced.

Third, I would like to point out a problem that will arise when Nunes' system is applied to a chain whose members are all phonetically empty. Consider a chain of a null subject in Italian as follows:

(27) *pro* ha telefonato.
he/she has-3sg telephoned
'He/she telephoned.'

(28) a. [~~he/she~~-CASE... [~~he/she~~-CASE...]]

b. [~~he/she~~-CASE... [~~he/she~~-CASE...]]

c. [~~he/she~~-CASE... [~~he/she~~-CASE...]]
=pro =pro

Assume that *pro* firstly has a phonetic form as a pronominal, which I represent as *he/she*. Following Nunes, Case would be deleted in a head position in a Case-assignment configuration with the matrix T (28a). Chain Reduction applies to a tail to reduce the number of constituents of the chain (28b). In addition, one more Chain Reduction applies to the head, resulting in (28c), in which none of the occurrences are phonetically realized.

The derivation of a chain of an empty subject is less economical than that of a chain of other kinds, based on Nunes. Compare a derivation of the null-subject chain with that of other

kinds:

- (29) a. [what-WH did you eat ~~what-WH~~]
b. [John-CASE was kissed ~~John-CASE~~]
c. [~~he/she-CASE~~ ha ~~he/she-CASE~~ telefonato]
 =pro =pro

In the *wh*-interrogative (29a) [u-wh] is deleted in a head of a chain; Chain Reduction has only to apply to a tail. In (29b) Case is deleted in a head position; Chain Reduction applies to a tail only once. In (29c) too Case is deleted in a head position. Chain Reduction, however, applies twice to both a head and a tail, which results in a *pro* chain. The empty subject chain is derived from a syntactic operation less economical than the *wh*-chain or the chain with the overt subject. The consequence sounds to be strange, with economy of articulation like *Avoid Pronoun Principle* (Chomsky 1981:65) taken into consideration: a derivational output (of a chain) without phonological realization should be more economical than that with phonetic materials.

Assume alternatively that formal features as a pronominal without phonetic materials, which I tentatively notate as $FF(pro^i)$,²⁰ are numerated, and let us account for a derivation following Nunes' argument of covert movement. Suppose the derivation has proceeded to a stage of merging T, resulting in [T [_{VP} FF(pro^i) ha telefonato]].²¹ Assume FF(pro^i) is probed by T and moves to [Spec,TP], resulting in [_{TP} FF(pro^i) T [_{VP} FF(pro^i) ha telefonato]]. It might be claimed that since a resulting chain of the formal features would consist of nondistinct copies (i.e. <FF(pro^i),FF(pro^i)>), the chain would not be formed, thus not be subject to Chain Reduction. [u-Case] is forgotten, though: the derivation will in effect result in [_{TP} FF(pro^i -CASE) T [_{VP} FF(pro^i -CASE) ha telefonato]], forming a chain <FF(pro^i -CASE),FF(pro^i -CASE)>. The chain composed of distinct copies must be subject to Chain Reduction; it is unclear how an account would continue under Nunes' system.²²

Fourth, it is uncertain whether it is ensured that uninterpretable features are in effect deleted in a head position of a chain as Nunes claims. The chain is a set of occurrences identical with each other; the uninterpretable features are deleted from the set of occurrences, namely from the chain itself (Chomsky 2000:116). Let us consider the previous example 'John was kissed.' We saw that Case deletion is represented as follows, following Nunes:

- (30) [John-CASE was kissed John-CASE]

Case is deleted in the head position in a Case-assignment relation with the matrix T (Nunes 2004:32). Actually, Case will be deleted in all or none of the positions, on the assumption that the uninterpretable features are deleted from the whole chain:

- (31) a. [John-CASE was kissed John-CASE]

b. [John-CASE was kissed John-CASE]

John's Case is not eliminated in any positions before Case deletion or in an unsuccessful deletion (31a); Case will be deleted in both a head and a tail in a successful case (31b). Namely, the problem is that assuming the identical nature of occurrences of a chain as well as uninterpretable feature deletion from the entire chain, it seems to be difficult to determine exactly at which position in a chain, either a head (, an intermediate position if any) or a tail, the uninterpretable features are deleted.

In sum, the mechanisms of phonological feature assignment to a chain introduced above do not seem to work well. Based on Chomsky (2004), internal Merge after S-O will be impossible in any way, with the identical nature of occurrences taken into account. Assuming Nunes (1999, 2004), derivations of in-situ S-O do not ensure uniformity of NS and/or SEM; derivations/NS outputs of a null subject chain are against economy of articulation. Further, it is impossible to determine exactly from which position in a chain uninterpretable features are deleted, on the assumption that occurrences are identical with each other and that the uninterpretable features are deleted from the entire chain.

2.2.3. A brief sum. Summarizing section 2, I have argued that the current architecture of phonological components – phonological features are registered in LEX together with syntactic/semantic features, introduced into an NS derivation with the latter, and stripped away from the derivation at S-O – should be improved from both the LEX and the PHON interfaces. On the LEX side, the assumption that idiosyncratic properties lie in LEX will not ensure uniformity of NS (and maybe SEM too) for all languages. On the PHON side, i) since timing of S-O and a stripping-away position of phonological features are determined both by convergent NS operations before S-O and by the principles which lead to a convergent NS component, presence/absence of uninterpretable phonological features will not affect NS-derivations; and ii) proposed mechanisms of phonological feature assignment to a chain will not work well: internal Merge after S-O seems to be difficult to be maintained unlike Chomsky (2004); in-situ S-O and a null subject chain are not sufficiently accounted for in terms of Nunes' (1999, 2004) system.

3. Late insertion of phonological features. Sigurðsson (2003) convincingly argues that language-particular property should exclusively be attributed to PHON. The fact that a language does not express a certain feature with a grammatical (i.e. physical) form does not mean that the feature is absent from the SEM of the language; for instance, the fact that Russian and Finnish do not have articles does not imply that they lack definiteness (Sigurðsson 2003:329). This means that all languages access all features of the universal set: language has innate SEMs independent of their physical realization (Sigurðsson 2003:333). Therefore, language variation is confined to PHON (Sigurðsson 2003:331). Sigurðsson is in line with my argument: I have argued that assumption that LEX is different among languages does not ensure uniformity of NS (and SEM). However, he does not present a syntactic model and a derivational mechanism which realize his claim; he does not clarify his position of how

LEX should be dealt with in a syntactic model, assuming language variation to be attributed to PHON, either. In this section I attempt to establish a model which ensures uniformity of NS and SEM, language-particular property lying in PHON. Further, I provide accounts for the issues that appear to be problematic based on the traditional view, which I pointed out in the last section, based on the new proposed model.

3.1. Distributed Morphology and proposal. Halle and Marantz (hereafter H&M, 1993) proposes a system called *Distributed Morphology* DM, a system of late insertion of phonological features. NS and SEM consist only of semantic/morphosyntactic features; the features are introduced into NS without phonological features (H&M 1993:121). After S-O an NS product is sent to *Morphological Structure*, where *Vocabulary insertion* takes place. Each Vocabulary entry of a language consists of two sets of features, phonological and semantic/morphosyntactic features. Vocabulary insertion finds an entry in which information of semantic/morphosyntactic features sent to Morphological Structure is matched with that of phonological features, and maps the phonological features of the entry onto the feature complex of corresponding semantic/morphosyntactic features. Categorical and subcategorical information can also come at the point of Vocabulary insertion (H&M 1993:122). Complexes of semantic/morphosyntactic features are not necessarily identical with those of actually occurring Vocabulary items of the language: ‘insertion requires only that a feature bundle of the Vocabulary item be nondistinct from features of a terminal node at M[orphological] S[tructure] that serves as a site of insertion’ (H&M 1993:121). Consequence is that with a structure of words determined by an NS operation (H&M 1993:113), linear order relation among morphemes is determined only at PHON; at the other levels there is a hierarchical relation only (H&M 1993:115).

H&M propose morphological operations, *merger* and *fusion*. Merger ‘joins terminal nodes under a category node of a head ... but maintains two independent terminal nodes under this category node,’ while fusion ‘takes two terminal nodes that are sisters under a single category node and fuses them into a single terminal node’ (H&M 1993:116). Based on the two operations, derivation of inflectional morpheme is accounted for as follows. See below:

(32) He ate the apple.

It is assumed (Chomsky 1995) that a finite verb does not raise to T in English. A syntactic representation is as follows:

(33) $[_{TP} \text{he } T_{[past]} [_{VP} \text{eat the apple}]]^{23}$

T merges with the finite verb under adjacency; then, fusion takes place, resulting in $[\dots[[\text{eat}] + T_{[past]}]\dots]$. A vocabulary that matches the information of $[\text{eat}] + T_{[past]}$, namely *ate*, is selected in Vocabulary insertion; corresponding phonological features are mapped onto the feature complex (H&M 1993:134-136).

DM appears to be an ideal system which realizes the claim that language variation is

exclusively confined to PHON (Sigurðsson 2003), since phonological features are introduced only after an NS derivation according to this system. I would like to raise points unclear to me, though. It is assumed that semantic/morphosyntactic features, which compose NS and SEM, ‘are more less freely formed’ (H&M 1993:121): it is not specified where those features come from. Thus, it seems to be difficult to specify where semantic/morphosyntactic features are located in a syntactic model: are all of them features at the universal level or at a language-particular level?; are they subject to Numeration?; are they all introduced at the start of a derivation or do some of them enter in the course of a derivation?; and so forth. Namely, the problem is that it is unclear how features referred to in DM interact with derivational mechanism(s). Therefore, I would like to claim that late insertion of phonological features should be organized within feature system of the current framework (Chomsky 2000~).

In addition, it is assumed that not only phonological features but also categorial features are inserted after NS. To construct a sentence, information of, say whether an item is N or V will have a crucial effect in an NS operation. See below:

- (34) a. Caesar destroyed the city.
 b. Caesar’s destruction of the city

The same argument structure of *destroy* is realized in both cases above, though a selecting head is V in (34a), but N in (34b). If *destroy* in (34a) had not yet been specified as V, T would not be merged. That is, if property of V had not yet appeared before Merge of T, selection relation between V^{24} and T would not be established; therefore, insertion of T too would not be ensured. If *destruction* in (34b) had not yet been specified as N, its Case assignment property would not be clarified either: property of (the Gen(itive)) Case assignment to *Caesar* and of *of*-insertion would remain to be unclear. Crucial is Case valuation. In Agree between V and a direct object, V assigns an Acc Case value to the latter (Chomsky 2000:123-124). If categorial information of V were not specified, [destroy] could not value a Case of [the city]: it is not (semantic feature) [destroy] but its categorial status as V that assigns an Acc Case value to [the city]. In the same way, if categorial information of N were not specified, [destruction] might assign a Case value to [the city]. That is, if valuation did not take place/inappropriately took place, the following PHON outputs would be predicted, contrary to the fact:

- (35) a. *Caesar destroyed of the city.
 b. *Caesar’s destruction the city

Therefore, information of categorial features should already be given at the beginning of/in the course of NS operations.

Taking the problems above into account, I would like to propose a derivational model which incorporates late insertion of phonological features that DM argues into the current

system (Chomsky 2000~). Feature components are introduced into a derivation as follows: {F} (the universal set of linguistic features) → [F] (a subset in an individual language drawn from {F}) → LEX (the lexicon in the language assembled from [F]) → LA (a lexical array accessed from LEX) → NS (Chomsky 2004:107). Assume i) that {F} contains as properties of human language features other than phonological ones, that is semantic, morphosyntactic, and categorial features; ii) that those features directly enter NS; and iii) that phonological features are introduced after NS, based on H&M (1993). That is, I would like to propose to eliminate the stages from [F] to LA from the model above, resulting in {F} → NS. One-time assembly from {F} in a language in the current model implies that [F] of the language may contain semantic/morphosyntactic/phonological features which are not contained in another language. If the stage like [F] which can generate a language property different from another does not exist in a model, specific forms that may be differed among languages cannot be generated in a later stage of a derivation either. If features in {F} universal for any languages directly enter NS, a contradiction between uniformity and feature selection from {F}, which Sigurðsson (2003) points out, is solved: assuming that {F} contains {F₁, F₂, F₃, F₄, F₅}, all of them, not some of them, are involved in an NS operation of any languages.

Eliminating the stages from [F] to LA, LEX must stand somewhere in a model. Assuming that nothing particular for an individual language is introduced in the course of a derivation, room for a level of LEX that deals with a language-particular property is left only after an NS operation. A candidate of such a stage is found in the DM system, namely Morphological Structure after S-O. In Morphological Structure Vocabulary insertion finds a Vocabulary entry which consists of two sets of phonological features and semantic/morphosyntactic features; the phonological features of the entry are then mapped onto a feature complex of the corresponding semantic/morphosyntactic features sent to Morphological Structure (H&M 1993:122). It is somewhat unclear, though, whether Morphological Structure and Vocabulary insertion accompanied by the former function not only as mapping of phonological features onto corresponding semantic/morphosyntactic features after S-O, but also as mental lexicon, stock of vocabularies. Here, I define LEX as a system unifying Morphological Structure and mental lexicon that works at PHON: LEX is not simply a list of exceptions, but also has a mechanism-like property which combines semantic/morphosyntactic features with phonological features. I summarize an outline of the proposed model:

- (36) A new computational model
 {F} (semantic/morphosyntactic/categorial features universal for all languages)
 ↓
 NS
 ↓
 PHON (, where LEX maps information sent from NS onto phonological features)

I assume that there exists only one interface with PHON in the model. The model shares property of a single output with the others (e.g. Groat and O’Neil 1996; Pesetsky 2000; Bobaljik 2003); I do not assume multiple S-O like Uriagereka (1999) and Chomsky (2000~).

The model here shares with Bobaljik (2003) property that a final NS output is sent to both PHON and SEM, though I do not refer to SEM in the model (36) concentrating on the PHON interface. Claiming that {F} directly enters NS, I suppose that NS operations like Merge and Move proceed only with features. Thus, I consider movement to be feature movement, which is rejected by Chomsky (2000). In addition, on the assumption that phonological features do not exist in an NS operation, I reject a notion of pied-piping.

From the formulation above, a case in which a concept of a referent is directly associated with a name will also be a result of syntax. For instance, \mathcal{D} is called *moon* in English. Imagine that the word *moon* comes out from inside ourselves when we see \mathcal{D} . It would not seem to be intuitively plausible that the word form *moon* comes out as a result of some syntactic operation; rather, the word would appear to come out directly from the stock of vocabularies. I deal with such a case in the following way: a semantic/categorial feature of [moon] + N in {F} goes straight to S-O without stopping at any intermediate stage of NS.

3.2.1. Pending issues. Hereafter, I discuss the problems raised in the previous section in turn, based on the syntactic model proposed above. I leave phonological feature assignment to a chain for the next section.

First, I pointed out that a model in which a derivation starts from LEX will not ensure uniformity of NS and SEM: in some cases different LAs can differ NS operations, though SEM are nondistinct; in other cases selection property of a head may differ NS derivations despite the same numeration. Here, I provide tentative accounts for the cases based on the proposed model. See below:

- (37) a. Il telefonerà.
 he telephone-FUT-3sg
 ‘He will telephone.’
- b. kare-ga denwasuru-darou.
 he-Nom telephone-FUT
 ‘He will telephone.’

(37a-b) is a case in which NS operations would differ based on different LAs, though SEM is nondistinct: only Italian (37a), but not Japanese (37b), would be argued to contain Agr, since Italian has agreement system.²⁵ Assume that T, Agr, (semantic features of) [he] and [telephone], N, V, and Case are selected from {F} to enter an NS derivation, and Merge takes place, resulting in [Agr [T [[[he]+N+Case][[telephone]+V]]]] in both cases.²⁶ Assume that the subject and the verb move, resulting in [_{AgrSP} [[he]+N+Case][[telephone]+V+T+Agr] [_{TP} t_{[telephone]+V+T} [_{vP} t_{[he]+N+Case} t_{[telephone]+V}]]], which is spelled out and sent to PHON, where LEX works.²⁷ Assume that after fusion takes place for [[telephone]+V+T+Agr], phonological features /telefone-/ are mapped onto [telephone], /-r-/ onto T, and /-à/ onto Agr in Italian, while phonological features /denwasuru/ are mapped onto [telephone], and /-darou/ onto T, but nothing onto Agr in Japanese. As a result, a PHON output of Italian is *il telefonerà*, while

that of Japanese is *kare-ga denwasuru-darou*. Alternatively, it will be assumed that Agr and Case, which do not affect an NS operation, are introduced after S-O, as H&M proposes (H&M 1993:135).²⁸ An output in the case is $[_{TP} [[he]+N] [[telephone]+V+T] [_{vP} t_{[he]+N} t_{[telephone]+V}]]]$ in both Italian and Japanese; Agr and Case are introduced in PHON for Italian, while Case, but not Agr, is introduced for Japanese; corresponding phonological features are mapped onto each feature. Either way, the NS output does not differ between Italian and Japanese. Since different LAs cannot be selected in the system, uniformity of NS (and SEM) is maintained with surface appearance among languages attributed to PHON operations.

See below, too:

- (38) a. *sa-ya-shtarii* *Zayd-un dar-an.*
 FUT(T)-3sg.MASC(Agr)-buy Zayd-Nom house-Acc
 ‘Zayd will buy a house.’
- b. *legge-va-no.*
 read-(T)-3pl.(Agr)
 ‘They read.’

(38a-b) is a case in which selection property of functional categories would be differed among languages: T selects Agr in Arabic (38a), while Agr selects T in Italian (38b)²⁹ (Ouhalla 1991:113). Assume that selection order of functional categories is [Agr[T[V]]]. Assume that in (38a) Agr, T, N, V, two Cases, (semantic features of) [buy], [Zayd], and [(a) house] are numerated and Merge takes place, resulting in [Agr [T [_{vP} [[Zayd]+N+Case][[[buy]+V][[a house]+Case]]]]]; in (38b) Agr, T, N, V, Case, [read], [they] are numerated and enter a derivation, resulting in [Agr [T [_{vP} [[they]+N+Case][[read]+V]]]]. Assume that the verb moves, resulting in [_{AgrP} [[buy]+V+T+Agr][_{TP} $t_{[buy]+V+T}$ [_{vP} [[Zayd]+N+Case][$t_{[buy]+V}$ [[a house]+Case]]]]] (38a) and in [_{AgrP} [[read]+V+T+Agr][_{TP} $t_{[read]+V+T}$ [_{vP} [[they]+N+Case] $t_{[read]+V}$]]] (38b) respectively.³⁰ Assume that after they are spelled out and sent to PHON, fusion takes place for the verbal complex [[buy]+V+T+Agr] (38a) and for that of [[read]+V+T+Agr] (38b); phonological features /shtarii/ are mapped onto [buy], /sa-/ onto T, and /-ya-/ onto Agr in the former, while phonological features /legge-/ are mapped onto [read], /-va-/ onto T, and /-no/ onto Agr in the latter. Notice that application of fusion may convert (hierarchical) order of categories built in NS into a different one after the operation. We saw above that [...T_[past][_{vP} [eat]...] is merged and fused as [...[[eat]+T_[past]]...]. In the same way, after fusion, the actual morphological order of [[buy]+V+T+Agr] is [T+Agr+[buy]+V] (*sa-ya-shtarii*) (38a), though that of [[read]+V+T+Agr] remains to be [[read]+V+T+Agr] (*legge-va-no*) (38b). With phonological features assigned to the other components,³¹ resulting PHON outputs are *sa-ya-shtarii Zayd-un dar-an* (38a) and *legge-va-no* (38b) respectively. Alternatively, assuming that Agr and Case do not exist in NS, they will be introduced only at PHON, as stated above. Adopting late insertion of phonological features and morphological operations, it is unnecessary to assume difference in selection property of functional categories among languages: uniformity of NS and SEM is strictly maintained, with

language-particular differences attributed to PHON. Recall that the accounts above are tentatively provided; in the next section I discuss how features like Agr and Case should be dealt with, how chains of null subject or head movement are formed, and so forth, modifying some of the accounts.

Second, I argued that timing of S-O and a stripping-away position of phonological features should be determined both by convergent NS operations before S-O and by the principles which lead to a convergent NS component. I repeat cases of the superiority effects below:

(39) a. Who said what?

b. *What did who say t_{what} ?

I claimed that presence/absence of uninterpretable phonological features will not affect NS derivations. The problem is now solved: assuming that phonological features enter after S-O, an NS component is generated with semantic/morphosyntactic/categorial features only, with the principles that lead to a convergent derivation applying to those features.

3.2.2. Final pending issue. As the final issue, I discuss phonological feature assignment to a chain, based on the proposed model. I argued in section 2.2.2. that mechanisms of phonological feature assignment to a chain proposed by Chomsky (2004) and Nunes (1999, 2004), which are based on the traditional view that phonological features are introduced at the beginning of a derivation and stripped away from an NS component at S-O, do not work well: internal Merge after S-O does not seem to be tenable in any way in Chomsky (2004); in-situ S-O and an empty subject chain are not convincingly accounted for in Nunes (1999, 2004). Another mechanism of phonological feature assignment to a chain must be found. Assuming late insertion of phonological features, no phonological features exist in a chain until an NS operation has finished. Then, the point will shift from which position in the chain phonological features are deleted from, to which position in the chain phonological features are assigned to in PHON.

A mechanism which would be applicable to the model proposed here is Groat and O'Neil's (1996) Single-Level Model. A computational character of pre-Spell-Out should not be different from that of post-Spell-Out in that they both operate on phrase-markers, create chains, and so forth; therefore, a uniform computational system is desirable (Groat and O'Neil 1996:119). Strong/weak features move before Spell-Out in all languages; a final phrase-marker of an NS derivation makes a PHON input. The difference in overt/covert movement depends on which position in a chain, a head position or a tail, is pronounced. A strong feature is like an affix, and it needs a host of a phonological material; therefore, the phonological feature is copied onto the head position (i.e. overt movement). If a functional category of a language has a weak feature, the phonological feature is not copied in accordance with the economy principle (Groat and O'Neil 1996:123-125). Based on the system, overt/covert *wh*-movement will be accounted for as follows. See below:

- (40) a. what did you eat t_{what} ?
- b. kimi-wa nani-o tabemashi-ta-ka?
 you-Top what-Acc eat -past-Q
 ‘what did you eat?’

Both *what* of English and *nani-o* of Japanese move to a higher position and form a chain with its copy, resulting in <what...what> (40a) and <nani-o,nani-o> (40b) respectively. Assuming that an English C has a strong *wh*-feature, phonological features are copied onto a higher position, resulting in overt movement; assuming, on the other hand, that a Japanese C has a weak *wh*-feature, phonological features are not copied, resulting in S-O in situ (Groat and O’Neil 1996:130-131).

Groat and O’Neil’s model would appear to be desirable, since uniformity of NS and SEM is maintained: on the assumption that all features, either strong or weak, move before S-O, phonological feature assignment to a chain is determined in terms of whether a language has a strong feature (S-O in a head position) or a weak feature (S-O in-situ). Their mechanism, however, is based on strong/weak feature, which component has already gone from the current system (Chomsky 2000~). Chomsky states as follows: ‘[a]ny such approach (like Groat and O’Neil) requires either new UG principles or language-specific rules to determine how the choice is made’ (Chomsky 2001:44,ft.25). Another inconvenience will concern nature of resulting chains. SEM of (40a-b) is not distinct. However, *wh*-occurrences in the chain <what,what> have both strong *wh*-features (40a), while those in the chain <nani-o,nani-o> have weak *wh*-features (40b). This is a situation in which NS outputs are differed, though SEM outputs are fortunately not distinct.

Occurrences in a chain are identical with each other; uninterpretable features are deleted from the entire chain (Chomsky 2000~). Thus, as I claimed in section 2.2.2., a *wh*-chain will not be <what-WH,what-WH> argued by Nunes (1999, 2004), in which [u-wh] is deleted from the higher *what*; rather, a *wh*-chain before [u-wh] deletion will be <what-WH,what-WH>, in which both *whats* still have [u-wh], while a chain after [u-wh] deletion will be <what-WH,what-WH>, in which [u-wh] is deleted from both *whats*. Assuming further that an NS component that still has uninterpretable features cannot be spelled out, I would like to argue that a legitimate chain to be spelled-out will be only <what-WH,what-WH>: the chain must not be spelled out unless uninterpretable features are deleted from all positions in the chain, either a head or a tail. The chain is uniform: no *wh*-occurrences have [u-wh]; therefore, they are identical. Consequence is that before S-O there is no factor that determines which position in a chain is pronounced or whether no phonological features are assigned to the chain. Then, I would like to assume that a factor/feature that determines them will enter/be added to the chain after S-O, namely, at PHON in the proposed model. It has been argued that several features are allowed to enter a derivation after S-O. Previously I mentioned H&M’s claim that Agr and Case, which do not affect NS and SEM, are introduced after S-O (H&M 1993:135). I follow H&M in assuming Agr to enter a derivation after S-O. This appears to contradict with assuming an Agree operation in NS. Here, I use Agr as an actual realization of

an agreement morpheme: I mean that Agree in an abstract sense takes place in NS, but an Agr feature enters a derivation after S-O as an actual morphological realization of Agree. Unlike H&M, I do not assume that Case is introduced after S-O, following the current system (Chomsky 2000~): assuming that Case is valued in Agree either as Nom by a finite T, or as Acc by a transitive v* (Chomsky 2000:123-124), Case must be introduced into a derivation from its start. Erteschik-Shir (2001) convincingly argues, based on analyses of Object Shift, that features like Foc(us) and Top(ic), which do affect SEM, can be introduced post-syntactically violating the Inclusiveness Condition (Erteschik-Shir 2001:71; see also Rizzi 2004). I assume that the features like Agr, Foc, and Top, which are introduced after S-O, are determinants of a pronounced position in a chain. In the following subsections I discuss phonological feature assignment to A-chain, A'-chain, and head chain in turn.

3.2.2.1. A-chain. First, I would like to consider phonological feature assignment to an A-chain. I repeat a relevant case below:

- (41) a. John was killed.
 b. [John was killed John]

I focus on a chain composed by two *Johns*, omitting the details. Considering the identical nature of occurrences, the chain at the point of S-O is as follows:³²

- (42) <[John]+N+CASE,[John]+N+CASE>

No phonological features are assigned to the chain at this point; the chain consists of semantic features [John], categorial features N, and Cases; Cases are eliminated from every position in the chain.

I would like to base my account on the following observation:

- (43) a. Gianni ha telefonato.
 Gianni has telephoned
 b. Ha telefonato Gianni.

The literatures have argued that a preverbal subject is topicalized/defocalized (43a), while a postverbal subject is focalized (43b) (e.g. Belletti 2001, among the others). Rizzi (2004) claims that a subject position is provided with special discourse properties ('quasi-topicality') (Rizzi 2004:3). Based on their claim, I assume that [Top] is assigned to a higher position of the chain <John,John> after S-O, which assigns phonological features to the higher *John*, resulting in (43a).³³

I turn to a null subject chain. See below:

(44) a. pro ha telefonato.
 he/she has-3sg telephoned
 ‘He/she telephoned.’

b. [he/she ha [he/she telefonato]]

I concentrate on a chain consisting of two *he/shes*. The chain at the point of S-O is as follows:

(45) <[he/she]+N+CASE,[he/she]+N+CASE>

No phonological features are assigned to any components at this point. I notate features as a pronominal, [person], [number], and [gender], as a feature complex [he/she]. Cases should be deleted from both higher and lower positions before S-O.

Lambrecht (1994) states that every sentence has information structure (Lambrecht 1994:6), and that a sentence must have focus, but may not have topic (Lambrecht 1994:206). He claims that ‘a referent which is new to a discourse necessarily involves focus status of the corresponding constituent’ (Lambrecht 1994:262). A reasoning from his claim is that a constituent which carries new information thus focus cannot be deleted from a sentence. This being applied to the empty subject case above, the subject which is made to be empty should not carry focus; rather, it is defocalized, thus (can be) deleted from the sentence. Therefore, it seems to be plausible to assume that [-Foc] enters all the positions in the chain (45) after S-O, which can let both positions be phonetically empty.

3.2.2.2. A'-chain. Next, I would like to consider phonological feature assignment to an A'-chain. I repeat relevant cases below:

(46) a. what did you eat?

b. [what-WH did you eat what-WH]

(47) a. You saw what?³⁴

b. [what-WH you saw what-WH]

(48) a. kimi-wa nani-o tabemashi-ta-ka?
 you-Top what-Acc eat -past-Q
 ‘what did you eat?’

b. [nani-o-WH kimi-wa [nani-o-WH tabemashi-ta-ka]]

In all the cases above no phonological features are assigned to any positions at S-O, as assumed so far. C with [u-Q] enters a matching relation with *wh* with [Q], which deletes

[u-Q]; *wh* moves to [Spec,CP], which is available for C's EPP, to delete its own [u-wh] (Chomsky 2000:128). The chains of *whats* and *nani-os* in (46-48) generated before S-O are as follows, with the identical nature of occurrences taken into account:³⁵

- (49) a. <what-~~WH~~,what-~~WH~~>
 b. <what-~~WH~~,what-~~WH~~>
 c. <nani-o-~~WH~~,nani-o-~~WH~~>

I consider first difference between (46-47), and second turn to (48).

It has long been claimed that *wh*-phrase in *wh*-interrogative carries focus (e.g. Culicover and Rochemont 1983, Lambrecht 1994). There is a factor, however, which has not absolutely led one to the conclusion: the fact that *wh*-phrase normally does not receive accent when *wh*-interrogative is used to ask new information that corresponds to the *wh*-part, though it is a standard claim that a new information part receives focus, as noticed by Rochemont (1978:55). In (46a) the accent does not fall on *what*, but on the last sentential component *eat*:

- (50) What did you EAT?³⁶

Lambrecht (1994) accounts for the fact as follows: in *wh*-interrogative a presuppositional structure (i.e. part of open proposition, namely part of the entire sentence minus a *wh*-expression, is pragmatically presupposed in the discourse, but not the *wh*-expression itself) is marked by the form and position of the *wh*-phrase; therefore, the accent does not need to mark the focus (Lambrecht 1994:285). The accent falls on *what* in (47a), though:

- (51) You saw WHAT?

Another observation concerns superiority effects in multiple *wh*-interrogative (Chomsky 1995, Richards 1997, Pesetsky 2000). I repeat previous examples below:

- (52) a. Who said what?
 b. *What did who say _{t_{what}}?

It is widely known that languages like Spanish allow the pattern (52b), and even in English *wh*ich-phrases avoid superiority effects:

- (53) ?Qué compró quién?
 what bought who
 'Who bought what?'

(54) a. Which person bought which book?

b. Which book did which person buy?

According to Pesetsky (1987, 2000), the *wh*-phrases above are *D(discourse)-linked*: when the answers to the questions are supposed to be drawn from a set of individuals previously introduced into the discourse or be salient to both speaker and hearer, superiority is avoidable (Pesetsky 2000:16). Rizzi (2004) associates the notion of D-linking with topicality (Rizzi 2004:18). He assumes that a D-linked *wh*-phrase is attracted by C with both [Q] and [Top] (Rizzi 2004:22,ft.4).

Important is Zubizarreta's (1998) observation of accent assignment to multiple *wh*-question. See below:

(55) Who said WHAT?

(56) a. Who knows what WHO bought?

b. *Who knows what who BOUGHT?
(Zubizarreta 1998:94-95)

(55) appears to display that accent simply falls on the last sentential element. Taking (56) into consideration, however, at least one of the *wh*-phrases must receive accent in multiple *wh*-interrogative, with accent assignment to another sentential component prohibited. Zubizarreta makes the same observations in D-linked *wh*-interrogative cases: accent must not fall on a sentential element other than a *wh*-phrase:³⁷

(57) a. ?Qué compró QUIÉN?³⁸
what bought who
'Who bought what?'

b. *Qué le compró quién a MARÍA?
what DAT-CL bought who to Maria
'Who bought what for Maria?'

(58) a. ?Qué bebida trajo QUIÉN?
which drink brought who
'Who brought which drink?'

b. *Qué bebida trajo quién a la FIESTA?
what drink brought who to the party
'Who brought which drink to the party?'
(Zubizarreta 1998:95)

Let us summarize the observations concerning (multiple) *wh*-interrogative. I repeat the examples below:

- (59) a. What did you EAT?
b. You saw WHAT?
c. Who said WHAT?
d. ?Qué compró QUIÉN?
 what bought who
 ‘Who bought what?’

Wh-phrase does not normally receive accent when it is used to ask new information corresponding to the *wh*-phrase (59a). A *wh*-phrase which is not interpreted as a question of new information but only as an echo question does receive accent (59b). In multiple *wh*-interrogative at least one of the *wh*-phrases must receive accent (59c). A D-linked *wh*-phrase too receives accent (59d). What can be drawn from the observations?

A sentence must have focus (e.g. Lambrecht 1994). It is somewhat controversial, however, whether a sentence can have multiple focus: some are affirmative (e.g. Selkirk 1984, Gussenhoven 1983), but others are negative (e.g. Lambrecht 1994, Zubizarreta 1998). Lambrecht (1994) convincingly argues that multiple focus is impossible, ‘since a given assertion cannot have more than one focus and since a given proposition cannot express more than one assertion’ (Lambrecht 1994:329). He supports the claim, raising two facts: i) a single sentence cannot be clefted twice; ii) a multiple *wh*-question cannot contain a cleft construction (Lambrecht 1994:329-330), as illustrated below respectively ((61a-b) are from French):

- (60) *It is YOUR foot that it is HE that treads on.
 (cf. He treads on your foot.)
- (61) a. C’est qui qui a mangé le fromage?
 it is who who has eaten the cheese
 (cf. Qui a mangé le fromage? (‘who ate the cheese?’))
- b. *C’est qui qui a mangé quoi?
 it is who who has eaten what
 (cf. Qui a mangé quoi? (‘who ate what?’))

Following Lambrecht, I assume that a sentence can have only one focus. Claiming in accordance with tradition that *wh*-part which is used to ask new information corresponding to that part carries focus, it is safely assumed that *what* in *what did you EAT?* (59a) carries focus.

If only one focus is allowed for sentence, focus should be assigned to only one of the *wh*-phrases in multiple *wh*-interrogative like *who said WHAT?* (59c). Based on the fact that *wh*-phrase does not normally receive accent when it is used to ask new information except an echo question case like *you saw WHAT?* (59b), it appears that the *wh*-phrase which does not receive accent is assigned focus, contrary to the actual auditory effect.³⁹ On the other hand, based on the observation that a D-linked *wh*-phrase, which is somewhat topic-like (Rizzi 2004), receives accent, a *wh*-phrase, which I would like to call morphologically specified focus, does not seem to exert focalizing effect when it is assigned accent. Therefore, I assume that *wh* which does not receive accent and maintains focalizing effect is assigned [Foc], while *wh* which receives accent thus does not exert focalizing effect is assigned [-Foc].^{40,41}

Let us turn to *wh*-chains again. I repeat the relevant cases below:

(62) a. What did you EAT?

b. You saw WHAT?

c. Who said WHAT?

d. Qué compró QUIÉN?
 what bought who
 ‘Who bought what?’

(63) a. <what-~~WH~~,what-~~WH~~>

b. <what-~~WH~~,what-~~WH~~>

c. <who-~~WH~~,who-~~WH~~>, <what-~~WH~~,what-~~WH~~>

d. <qué-~~WH~~,qué-~~WH~~>, <quién-~~WH~~,quién-~~WH~~>

(62c-d) have a chain for each *wh*, resulting in two *wh*-chains (63c-d). Assuming the identical nature of occurrences, all the [u-wh]s are deleted from the chains at S-O (63). Lots of literatures have argued that in languages like English a *wh*- or focus phrase which moves to a peripheral position is interpreted as having quantificational force in the position (e.g. Chomsky 1995, Rizzi 1997). Therefore, I assume that a pronounced position in a *wh*-chain of the languages is determined in terms of whether [Foc] or [-Foc] enters a higher position in the chain. I suppose that when [Foc] is introduced, the higher position is assigned phonological features; when [-Foc] enters, the higher position is prevented from being assigned phonological features with the latter assigned to the lower position instead. *What* in (62a) is used to ask new information without receiving accent; [Foc] enters a higher position of the chain, which lets phonological features be assigned to the higher *what*. *What* in (62b), which is only interpreted as an echo question, does not exert focalizing effect receiving accent;

[-Foc] enters a higher position, which prevents the higher *what* from being assigned phonological features and causes the lower *what* to be assigned phonological features. In (62c) *who* does not receive accent, maintaining focalizing effect; [Foc] is introduced into a higher position of the *who*-chain, which lets the higher *who* have phonological features. *What* receives accent, losing focalizing effect; [-Foc] enters a higher position of the *what*-chain, which assigns phonological features to the lower *what*. (62d) is a case which is not subject to the superiority effects because of D-linking. *Qué* does not receive accent, keeping focalizing effect; [Foc] enters a higher position of the *qué*-chain, which lets the higher *qué* be assigned phonological features. *Quién* receives accent, losing focalizing effect; [-Foc] is introduced into a higher position of the *quién*-chain, which causes the lower *quién* to have phonological features.^{42,43}

I turn to a *wh*-in-situ case. I repeat a Japanese example and a chain representation below:

- (64) a. *kimi-wa nani-o tabemashi-ta-ka?*
 you-Top what-Acc eat -past-Q
 ‘what did you eat?’
- b. [nani-o-WH *kimi-wa* [nani-o-WH *tabemashi-ta-ka*]]
- c. <nani-o-~~WH~~,nani-o-~~WH~~>

Taking into account the identical nature of occurrences, all the [u-wh]s should be eliminated from the chain before S-O.

I would like to present two ways of analyzing *wh*-in-situ. One way is to simply assume that [Foc] enters a lower position of a *wh*-chain in languages like Japanese, which enables phonological features to be inserted into a lower *wh*. This is supported by the fact that in-situ focus is allowed in Japanese. See below:

- (65) *Watashi-wa ANO RINGO-GA/O tabe-tai -nda (kono ringo dewanaku).*
 I-Top that apple-Nom/-Acc eat-want-emphasizer (this apple not)
 ‘I want to eat THAT APPLE (, not this apple).’

Ano ringo-ga/o is interpreted as (contrastively) focalized in the in-situ position.

The other way will be one based on morphological property of a Japanese *wh*-phrase. Hagstrom (1998) proposes to analyze a Japanese Q-particle *-ka* as base-generated with a *wh*-part. His claim is based on the observation that a *wh*-part *nani* combines with a Q-morpheme *-ka* to become an existential quantifier *nani-ka* ‘something’, which is extended to the other kinds of *wh*-phrases. *-Ka* is a Q-head variable which has existential quantificational force (Hagstrom 1998:134); it takes a *wh*-part (the restriction set) as its argument (Hagstrom 1998:129), and yields a value which corresponds to the property of being a person (*choice function*, Hagstrom 1998:130). *-Ka* moves to a C head position in *wh*-interrogative (Hagstrom 1998:92).⁴⁴ Thus, based on Hagstrom, (64a) is derived in the following way:

- (66) a. [kimi-wa nani-ka-o tabemashi-ta]
 what-Q
- b. [kimi-wa nani-t_{ka}-o tabemashi-ta-ka]
 what Q

The Q-particle *-ka* is base-generated with the *wh*-phrase *nani* (66a); the Q-particle then moves to C, leaving a copy (66b).

Hagstrom's analyses indicate that in Japanese *wh*-part in effect does not move, but only Q-particle moves: in the construction above [...nani-t_{ka}-o...-ka] the *wh*-part *nani* remains in situ, but only the Q-morpheme *-ka* moves to C. At this point he turns to Watanabe's (1992) *wh*-in-situ account in terms of null operator movement: a phonologically empty operator moves to [Spec,CP] in overt syntax in Japanese *wh*-interrogative (Watanabe 1992:264). Hagstrom claims that what is moved is not a null operator as assumed in Watanabe (1992), but a Q-particle of a head category (Hagstrom 1998:92).

Chomsky (2000) assumes that C possesses [u-Q], which is deleted by the interpretable counterpart [Q] that *wh* has; *wh* also has [u-wh] and moves until it is deleted (Chomsky 2000:128). Hagstrom's claim for *wh*-in-situ is formulated in the way that '[wh] pied-pipes only the head (overtly or covertly)' (Chomsky 2000:128). Following him, a resulting chain in (66) will be as follows: <nani-t_{ka}-o,ka+[(u-)wh]>. It seems to be unclear whether the chain is legitimate: since the category in situ is a maximal projection while the category in the higher position is a head, the chain will not be uniform. Let us assume that [u-Q] is deleted by *wh*'s [Q] in English. Assume further that [Q] lies in Q-particle and [u-Q] is deleted by Q-particle's [Q] in Japanese (cf. Miyagawa 2001). That is, assume that Japanese separately has *wh*-chain and Q-chain (i.e. [nani-o-~~WH~~...nani-o-~~WH~~] and [ka-Q...ka-Q]), while English has only *wh*-chain into which [Q] is incorporated as in [what-~~WH~~-Q...what-~~WH~~-Q]. I stated that in languages like English a *wh*- or focus phrase exerts quantificational force in a peripheral position, which has led me to suppose that [Foc] enters a higher position in English. If Q-particle exerts quantificational force in *wh*-in-situ languages as Hagstrom claims, the situation appears to be the same as the languages like English in that a category with quantificational force always exerts the force in a peripheral position. Q-particle cannot have quantificational force in the original position, as illustrated below:

- (67) a. [kimi-wa nani-t_{ka}-o tabemashi-ta-ka]
 what Q
- b. [kimi-wa nani-ka-o tabemashi-ta]
 what-Q

(67b), in which the Q-particle remains in the original position, is not interpreted as *wh*-interrogative, but as 'you ate something.' Then, assume that [Foc] enters a higher position in Japanese Q-chain, which lets phonological features be assigned to the higher *-ka*. Suppose

further that [-Foc] enters a higher position in Japanese *wh*-chain since a sentence cannot have more than one focus (Lambrecht 1994); phonological features are then assigned to the lower *nani-o*.

Taking into account morphological differences among languages, the accounts based on Hagstrom appear to be desirable. It is maintained among languages that quantificational force is exerted in a higher position. The accounts will not ensure uniformity of NS, though: to assume that Japanese has both *wh*-chain and Q-chain as well as that English has only *wh*-chain equals to admitting that NS has two different derivations, one for Japanese and the other for English, though SEM outputs are not distinct. Further elaboration will be necessary to pursue this theoretical line.

3.2.2.3. Head chain. Finally, I turn to a head chain. I attempt to account for phonological feature assignment to head chain, incorporating head movement into NS, unlike Chomsky, who treats head movement as a phonological operation (Chomsky 2001:37). See below:

(68) a. John often kisses Mary.

b. *John kisses often Mary.

(69) a. Jean embrasse souvent Marie.

John kisses often Mary

b. *Jean souvent embrasse Marie.

The adverb *often* can precede, but cannot follow, the finite verb *kisses* in English (68); the adverb *souvent*, on the other hand, cannot precede, but can follow, the finite verb *embrasse* in French (69). This has standardly been accounted for in the way that an English finite verb cannot, but a French finite verb can, move to (v-to-)T (Emonds 1978, Pollock 1989, Chomsky 1995, among others). Several peculiarities of head movement have been presented; for instance, i) head movement is subject to more strict locality than phrasal movement (the *Head Movement Constraint* HMC, Travis 1984); ii) a moved head does not c-command its trace (i.e. in $[_{TP} [_T V+T]][_{VP} t_V]$ the moved V does not c-command the trace); and iii) head movement does not extend a tree, which violates the Extension Condition (Chomsky 2000). Head movement is now supposed to be an operation at PHON (Chomsky 2001:37). There are at least two reasons which seem to prevent one from assuming verb movement to be an NS operation. One is that verb movement does not have a semantic effect: whether the finite verb moves crossing over the adverb in French (69a) or not in English (68a), verb movement does not produce semantic difference between them. The other is that assuming that movement is in general triggered by some uninterpretable features (Chomsky 2000~), no uninterpretable features appear to be involved in verb movement: unlike a noun phrase and a *wh*-phrase, which are supposed to have [u-Case] and [u-wh] respectively, a moved verb itself does not seem to possess any uninterpretable features. Verb movement appears to be a movement only

for a morphological reason: T needs a host of phonological matrix on which it is realized (H&M 1993:137-138).

Several counterarguments against the account of verb movement in terms of PHON operation have been presented, though. Based on Chomsky (2001), NS outputs of (68-69a) are as follows:⁴⁵

- (70) a. [TP Jean T [_{v*P} souvent [_{v*P} t_{Jean} v [_{VP} embrasse Marie]]]]
 b. [TP John T [_{v*P} often [_{v*P} t_{John} v [_{VP} kisses Mary]]]]

The subject moves to [Spec,TP], leaving a trace; the finite verb moves neither in French nor in English, remaining in the original position. A PHON operation applies only to French, which results in the following PHON outputs:

- (71) a. [TP Jean embrasse+v+T [_{v*P} souvent [_{v*P} t_{Jean} t_{embrasse+v} [_{VP} t_{embrasse} Marie]]]]
 b. [TP John T [_{v*P} often [_{v*P} t_{John} v [_{VP} kisses Mary]]]]

The finite verb *embrasse* of French, but not the finite verb *kisses* of English, moves to T at PHON. The account would appear to be desirable in that uniformity of NS and SEM is completely ensured with difference between the languages attributed to the PHON operation.

Matushansky (2004), however, convincingly argues that verb movement at PHON is untenable. Assuming that H's complement β in [_{HP} α [H β]] is spelled out at the level of the phase HP (Chomsky 2000), it is impossible that in a configuration [... [_{v(*)P} v(*)] ... [_{VP} V ...]] (a French) V moves to v(*) after a complement of v(*) (i.e. VP) is spelled out (Matushansky 2004:10). Consequently, she claims that verb movement cannot be dealt with as an operation in PHON.

Holmberg (1999), Erteschik-Shir (2001), and Nilsen (2003) claim that verb movement does affect interpretation, against Chomsky (2001): (optional) verb movement is focus-driven, following them. Holmberg's Generalization (Holmberg 1986) states that when a weak pronoun is shifted (*Object Shift*), verb movement surely follows the pronominal movement. Erteschik-Shir shows that a moved verb in Object Shift is stressed, which displays that the moved verb is focalized (Erteschik-Shir 2001:50-51).

Based on the arguments, I would like to seek the way of claiming that verb movement is an NS operation. I stated that two reasons for which one is reluctant to assume verb movement to be derived in NS are i) that verb movement does not yield semantic effects and ii) that verb movement does not contain any uninterpretable features. The first point seems to be surmountable, based on the arguments by Holmberg (1999), Erteschik-Shir (2001), and Nilsen (2003): verb movement does affect semantic interpretation. The second point is translated into a problem of trigger of verb movement, namely a problem that verb movement does not appear to contain any trigger. The literatures have proposed a solution, counting on *c(ategorial)-selection* (Svenonius 1994, Julien 2002, and Matushansky 2004, among others).

Their basic claim is that c-selection is done by a feature and satisfied by head movement. Matushansky argues that c-selection in head movement corresponds to Agree in phrasal movement: c-feature of T selects V, while ϕ -features of T agree with matching features of a phrase. C-feature is formulated as [u-V], which is counterpart of V (Matushansky 2004:19; also Julien 2002).

Assuming that head movement is triggered by T's c-selection property, that is by [u-V], it follows that head movement must be done in NS: [u-V] must be deleted before S-O. Matushansky (2004) proposes that head movement is not different from phrasal movement, arguing not only that c-selection in head movement parallels Agree in phrasal movement but also that a head and a phrase both move and are adjoined to a root to extend a tree. Namely, whether Y or WP moves in a configuration [_{XP} X [_{YP} ZP [Y WP]]], the landing site is [Spec,XP]; Y movement results in [_{XP} Y X [_{YP} ZP [_{t_Y} WP]]], while WP movement results in [_{XP} WP X [_{YP} ZP [Y _{t_{WP}}]]]. Only in head movement does morphological merger (*m-merger*), a head-merging operation, take place, resulting in [_{XP} [_X Y X] [_{YP} ZP [_{t_Y} WP]]] (Matushansky 2004:23-24).

The account of verb movement by Matushansky would appear to be mostly desirable. Assuming that verb movement is triggered by [u-V] of T, it is incorporated into NS; uniformity of NS is maintained, since verb cross-linguistically moves to T to delete [u-V]. However, the account is silent for covert movement, as she puts aside difference in overt and covert movement (Matushansky 2004:25,ft.23). It seems that she bases the account on the assumption that phonological features are introduced at the beginning of a derivation. She suggests only that overt head movement is like Agree with pied-piping, while covert head movement is Agree without pied-piping, as in phrasal movement (Matushansky 2004:25,ft.23). To assume both kinds of head movement will prevent uniformity of NS from being maintained.

Julien (2002) claims, assuming late insertion of phonological features, that a finite V position on which T is realized is a T position, with V cross-linguistically moving to T. Namely, the position which is occupied by *kisses* in (68a) and that which is occupied by *embrasse* in (69a) are both T positions. No matter how different a position which a finite verb occupies appears to be among languages, the position which the verb is required to move to does not differ: 'the finite verb is marked for only one of the inflectional categories that ... necessarily represent sentential heads [, which] is tense.' 'It is possible ... that the finite verb in English [, and so forth,] does indeed acquire the tense suffix by moving to a temporal head' (Julien 2002:272-273).

Given that V cross-linguistically moves to T, the account of verb movement by Julien would appear to be optimal, since not only uniformity of NS and SEM but also that of PHON are strictly maintained among languages: the position to which a finite verb moves is exactly the position at which the verb is spelled out and assigned phonological features. Julien's account, however, will not be without problem. Her claim is based on Cinque's (1999) adverb hierarchy. See below:

(72) a. John completely forgot her instructions.

- b. John forgot her instructions completely.

The adverb *completely* can appear in different positions, either before a finite verb (72a) or at the end of a sentence (72b). (72a) means that John forgot her instructions at a certain time; (72b), on the other hand, is ambiguous in that it means not only the reading in (72a) but also that any part of her instructions disappeared from John's mind. According to Cinque, an adverb occupies a Spec of a projection headed by a specific feature: *completely* occupies a Spec of a projection headed by *completive*. The fact that *completely* produces different semantic effects in each position is interpreted in the way that there is a projection for each *completive*, thus a Spec for each *completely*. Adverb order is represented as follows (I omit the details):

(73) [...[completive_{1P} completely ... [...[completive_{2P} completely

It can be seen from (72a-b) that the finite verb *forgot* is situated between the two projections. Therefore, it is determined that an English finite verb occupies a position between those projections (Julien 2002:271).⁴⁶

Julien notices that TP is higher than the *completive_{1P}* in Cinque's hierarchy (Cinque 1999:106), as illustrated in

(74) [...[TP T ... [...[completive_{1P} completely ... [...[completive_{2P} completely

This would mean that the English finite verb cannot move to T. Julien claims, though, that there appear to be no adverbials which are surely located in the Spec of TP. She raises *now* and *then*, which Cinque argues occupy [Spec,TP]. It would be predicted that the adverbs must be located only in left-peripheral positions, contrary to the facts:

- (75) a. Now Mary smokes cigars.
 b. Mary now smokes cigars.
 c. *Mary smokes now cigars.
 d. Mary smokes cigars now.
- (76) a. Then Mary smoked cigars.
 b. Mary then smoked cigars.
 c. *Mary smoked then cigars.
 d. Mary smoked cigars then.

Now and *then* can appear not only in the left-peripheral positions (75-76a) but also after the subject (75-76b) and clause-finally (75-76d). This indicates, Julien argues, that T(P)'s position cannot be clarified in terms of adverbial order. Therefore, she concludes that the position of T can be determined only by seeing the position of the finite verb on which tense is realized (Julien 2002:272).

Shortcomings of Julien's account seem to lie in her methodology: though T's position is in effect not specified in a sentence, the position is cross-linguistically assumed to be a position which a finite verb occupies. Thus, based on her argument, neither will T's position be testified; nor will possibility be denied that the finite verb is spelled out in a position other than T, plausibly in the original position V.

I summarize the problems of the analyses that attempt to incorporate head movement into NS. Following Matushansky (2004), uniformity of NS will not be maintained, with both overt and covert verb movement assumed. Based on Julien (2002), it will not be ensured that a finite verb is always spelled out in T, since T's position would not be testified.

I would like to present a solution to the problems above, considering verb movement in terms of phonological feature assignment to a head chain. I agree with Julien (2002) in supposing that phonological features are introduced only after S-O, as done so far. Following Matushansky (2004) (and Julien (2002)), I assume that trigger of verb movement is [u-V], a c-selection feature, of T. An outline is as follows: parallel with chains of A-/A'-movement, I assume that categorial V-feature, to which verbal semantic feature merges, cross-linguistically moves in NS, deleting T's [u-V]; a feature complex composed of the V- and its semantic features forms a chain with its copy; phonological features are assigned to a position in a chain by some determinant feature which is introduced in PHON. I repeat relevant cases below:

(77) a. Jean embrasse souvent Marie.

b. John often kisses Mary.

An NS output and a resulting chain are as follows:⁴⁷

(78) a. [TP ...[kiss]+V+v*+T ...[_v*P ... t_{[kiss]+V+v*} [VP t_{[kiss]+V} ...]]]

b. <[kiss]-V-v*-T, [kiss]-V-v*, [kiss]-V>

Notice that head chain seems to have property different from the other kinds of chains in that occurrences do not appear to be identical: the feature complex [kiss]+V raises, picking up a feature at each movement, v* in the first raising and T in the second one. There have been proposed a lot of derivational ways which attempt to overcome peculiarity of head movement, among which is Matushansky's (2004) account in terms of *m-merge* I introduced before. Matushansky in effect rejects the uniformity condition of chains in that it is a stipulation; she attempts to derive the nature of all kinds of movement chains from a series of operations,

Agree, Copy, and (Re)merge (Matushansky 2004:5-6).⁴⁸

Another way is assume an operation called *interarboreal operation* (Bobaljik and Brown 1997)/*sideward movement* (Hornstein 2001, Nunes 2004). The idea originates in elimination of D-Structure, which has traditionally been assumed in generative grammar (Chomsky 1981), and in transfer to a derivational system in terms of Merge (Chomsky 1995). A standard assumption has been that a category must move and merge to a syntactic object which contains an original copy of the category (e.g. Chomsky 1995): maintaining (a single-rooted) D-Structure, the moved category would have to be associated with its copy by being merged to the root. Deleting D-Structure from the system, there is no reason to maintain the assumption: the moved category will have an option of merging either with the root or with a newly numerated category (Nunes 2004:5-6). Head movement in terms of sideward operation proceeds as follows. Assume that after VP is generated, T is numerated⁴⁹; V in [_{VP} ... V ...] leaves a copy and merges with T, resulting in two syntactic objects, [_T V T] and [_{VP} ... t_V ...]; they merge, resulting in a larger syntactic object [_{TP} [_T V T][_{VP} ... t_V ...]] (Brown and Bobaljik 1997).

Those who advocate sideward movement seem to have to abandon the c-command condition, claiming that the Extension Condition should be superior to the c-command condition (Brown and Bobaljik 1997:352-353, Hornstein 2001:74). In [_{TP} [_T V T][_{VP} ... t_V ...]] the moved V does not c-command its copy. The resulting component does satisfy the Extension Condition, though, which leads them to assume the derivation to be legitimate.

The account by Nunes (2004), in which the c-command condition is still maintained, appears to be even contradictory inside his theory. His account in terms of sideward movement counts on *Conditions on Form Chain* by Chomsky (1995), which states that in a chain $\langle \alpha, \beta \rangle$, α c-commands β (Chomsky 1995:356, Nunes 2004:91). Phrasal movement is accounted for as follows. Assume that there are two syntactic objects, [_K ... α ...] and [_L ...]. Assume that α makes a copy and it merges to the other object, resulting in [_K ... α_i ...] and [_M α_i [_L ...]] respectively. Two α s do not form a chain, since they do not c-command each other. Assume that a head H is numerated, and takes [_M α_i [_L ...]] as its complement and [_K ... α_i ...] as its Spec, resulting in [_{HP} [_K ... α_i ...][_{H'} H [_M α_i [_L ...]]]]. Here too, no chain is composed because of absence of c-command relation between α s. Assume further that a head Y which is newly introduced in a later derivation requires α 's copy, resulting in [_{YP} α_i [_{Y'} Y ... [_{HP} [_K ... α_i ...][_{H'} H [_M α_i [_L ...]]]]]]. It is argued that the highest α can in principle form a chain. Based on the assumption that the lower α s are not in a c-command relation, it is supposed that the highest α forms two distinct chains, one with the α in [Spec,HP] and the other with the α in the complement of H. Then Chain Reduction takes place and reduces the lower α s in each chain, resulting in [_{YP} α_i [_{Y'} Y ... [_{HP} [_K ... α_i ...][_{H'} H [_M α_i [_L ...]]]]]] (Nunes 2004:94-95). The way of derivation in terms of sideward movement is extended to head movement. The derivation proceeds in the same way as Bobaljik and Brown's (1997) interarboreal operation, which I repeat next: i) generation of [_{VP} ... V ...] and numeration of T; ii) copy of V and its merge to T, resulting in [_T V_i T] and [_{VP} ... V_i ...]; iii) they merge, resulting in [_{TP} [_T V_i T][_{VP} ... V_i ...]]. In addition, it is assumed iv) that two V_is form a chain $\langle V_i, V_i \rangle$; and v) that Chain Reduction deletes the lower copy, resulting in [_{TP} [_T V_i T][_{VP} ... \forall_i ...]] (Nunes 2004:151).

Following the analyses of phrasal movement, however, two V_i s could not form a chain, since they are not in a c-command relation. This seems to be a contradiction in Nunes' account.

Covert head movement too does not appear to be convincingly accounted for. I repeat Nunes' (2004) account in terms of sideward movement of formal features introduced previously. After VP is generated, V's formal features sideward-move and adjoin to T, resulting in two syntactic objects $[_{VP} \dots V^i \dots]$ and $[_T FF(V^i)+T]$; they merge, resulting in $[_{TP} [_T FF(V^i)+T] [_{VP} \dots V^i \dots]]$; a resulting chain of V's formal features would be $\langle FF(V^i), FF(V^i) \rangle$. Since the chain consists of nondistinct copies, they cannot form a chain; therefore, they are not subject to Chain Reduction (Nunes 2004:153-154). Since the moved V itself does not have any uninterpretable features unlike a moved category in phrasal movement, the copies in the resulting chain would be argued to be nondistinct. Consider the fact that T is realized on V. T appears to be added to V at Merge of $FF(V^i)$ to T, though nothing extra is attached in phrasal movement: a resulting chain seems to be $\langle FF(V^i)+T, FF(V^i) \rangle$, in which the higher copy has property different from the original $FF(V^i)$. As long as the higher copy appears to be distinct from the lower one, it could not definitely be said that the copies do not form a chain, thus are not to subject Chain Reduction, unlike Nunes' claim.

In sum, all the accounts introduced above could be argued at the expense of some theoretical principle: Matushansky (2004) rejects the uniformity condition of chains; those who argue sideward movement demote the c-command relation, claiming that the Extension Condition is primary. What I would like to pursue is to maintain all those principles, while accepting that head chain has property different from the other kinds of chains. The notions of uniformity and c-command not only concerns chain formation but also the entire system: for instance, NS is assumed to be uniform; binding relations are defined in terms of c-command, and so forth. In my opinion, the principles must be maintained; rather, it is head chain that should be accepted as somewhat exceptional thus be provided a way of incorporating it into the whole system. I repeat the example of head chain below:

(79) $\langle [kiss]-V-v^*-T, [kiss]-V-v^*, [kiss]-V \rangle$

I pointed out that occurrences of head chain do not appear to be identical with each other: the feature unit $[kiss]+V$ moves, picking up a feature at each raising, v^* and T in turn. I would like to argue that this is source that differentiates head movement from the other kinds in both the way of derivation and the nature of the resulting chain. Namely, head movement surely adds a feature in each movement,⁵⁰ while phrasal movement does not; therefore, occurrences of head chain do not appear to be identical since the higher occurrence is added more features than the lower one, while occurrences of phrasal movement are added no features but their uninterpretable features are only deleted. I would like to attempt to solve the problem on head chain, counting on the notion of *extended projection* in Grimshaw's (1991) sense: C selects T, and T selects v^*/V ; in other words, TP and CP are extended projections of v^*/V . I would like to assume that as long as a projection is extended from a lower head, features which are added when the head moves are shared by all the copies of the head. Specifically, in the following chain

(80) <[kiss]-V-v*-T, [kiss]-V-v*, [kiss]-V>

I suppose that the copies in T, v*, and V share all the features after the feature complex [kiss]+V has moved to (v* to) T, resulting in the chain like below:

(81) <[kiss]-V-v*-T, [kiss]-V-v*(-T), [kiss]-V(-v*-T)>

The resulting chain as an NS output is, I assume, uniform in that all the copies of the moved head share the features which are added in each movement operation; therefore, the occurrences of the head are identical at S-O.

A position in the head chain to which phonological features are assigned must be determined. I have argued that a pronounced position should be decided in terms of some feature which enters a chain after S-O. Previously I referred to H&M's claim that a feature which does not affect NS and SEM, is introduced after S-O, an example of which is Agr (H&M 1993:135). Following them, I assumed Agr to enter a derivation after S-O as an actual realization of Agree. Julien (2002) argues that Agr is added to already existing syntactic heads: being not syntactic heads by themselves, Agrs can be scattered into different heads (Julien 2002:254). I would like to incorporate Julien's claim into the theory here, and suppose that the feature [Agr], which is introduced into a head chain after S-O, determines a pronounced position in the head chain. It has been argued in the literatures that a language in which a finite verb is situated in a higher position has rich agreement system, while a language in which a finite verb is located in a lower position has only poor agreement morphemes (Chomsky 1981, Rizzi 1982, Roberts 1993, among others). Therefore, I assume that [Agr] enters a higher position of head chain in languages like French with rich agreement system, which makes it possible for phonological features to be assigned to the higher copy, resulting in

(82) Jean embrasse souvent Marie.

As for languages with poor agreement morphemes, I assume that [Agr] is introduced into a lower position, which lets the lower copy have phonological features, resulting in

(83) John often kisses Mary.

4. Conclusion. In this paper I compared the traditional view on phonological features, namely the view that phonological forms are registered in LEX, are introduced into a syntactic derivation, and stripped away from the syntactic object at S-O (e.g. Chomsky 1995, 2000~) with late insertion of phonological features, which assumes that phonological features are introduced by morphological operations only after an NS operation (DM, H&M 1993). Based on the notion of uniformity, I firstly claimed that assuming idiosyncracies to lie in LEX of each language does not maintain uniformity of NS and/or SEM. Next, I discussed problems that are raised on the PHON interface. I argued the followings: i) presence/absence of

(uninterpretable) phonological features will not be related to whether NS derivations converge or crash, unlike Chomsky (2000~); and ii) proposed mechanisms of phonological feature assignment to a chain do not seem to work well in that internal Merge after S-O seems to be impossible (Chomsky 2004) and that both in-situ S-O and an empty subject chain do not appear to be convincingly accounted for (Nunes 1999, 2004). I introduced DM and pointed out that since origin of features referred to in DM is somewhat unclear, interaction between the other derivational systems and the features is also unclear. I attempted to organize late insertion of phonological features within the current system, proposing a model in which features in {F} directly enter NS and LEX works at PHON, eliminating intermediate stages assumed in Chomsky (2000~): {F} → ~~[F]~~ → ~~LEX~~ → ~~LA~~ → NS → PHON (, where LEX works). I turned again to the problems that are raised from the traditional view above. I claimed that strict uniformity of NS and SEM is maintained, with all surface differences which apparently belong to property of individual lexical items attributed to operations at PHON. I discussed phonological feature assignment to a chain in detail. Assuming uniformity of a chain and identical nature of occurrences, I argued that a chain must not be spelled out unless uninterpretable features are deleted from all the occurrences of the chain. Based on the literatures which claim that features like [Agr], [Foc], or [Top] can be introduced after S-O (e.g. H&M 1993, Erteschik-Shir 2001, Rizzi 2004), I proposed that one of the features can enter a chain after S-O and determines a position which is actually pronounced in the chain. In this way I attempted to provide accounts for the issues that are problematic from the traditional view, based on the proposed model.

Notes

* Part of this paper was presented at the 20th Scandinavian Conference of Linguistics, Helsinki, January 2004, and the Antisymmetry Approaches at the Interfaces Conference, Montréal, October 2003. I am thankful to the audiences at these presentations for comments and suggestion. Special thanks to Anders Holmberg and Juan Uriagereka for their helpful comments, and to Fred Karlsson for his heartfelt advice in editing the style. I alone am responsible for any errors.

¹ Hereafter, I notate Chomsky's current system (i.e. Chomsky 2000, 2001, 2004) as *Chomsky (2000~)*, except when particular reference to each paper is necessary.

² PHON and SEM have traditionally been called PF and LF respectively. Hereafter, I unify terminology into PHON and SEM both for the current system since Chomsky (2000~) and for the traditional system until Chomsky (1995), except when it is necessary to make particular references to them.

³ I will return to phases later.

⁴ T is assumed to be defective, if selected by V (Chomsky 2000:102).

⁵ An asterisk * is adjoined when *v* has full argument structure, namely transitive *v* or experiencer (Chomsky 2001:43,ft.8).

⁶ I turn to this property later.

⁷ Pesetsky and Torrego (2001:359) explains the EPP property more in detail than Chomsky: 'EPP is a property of a *feature* of a head – not a property of the head itself. Thus, a head that bears features F and G might have the EPP property for F, but not for G. EPP is thus, a "subfeature of a feature."'

⁸ Φ -PHON is said to highly differ among Ls, which I turn to later.

⁹ Agr has entirely been eliminated from the system since Chomsky (1995). Later, I assume that features like Agr which do not affect NS and SEM may enter after S-O, following Halle and Marantz (1993) and Erteschik-Shir (2001).

¹⁰ I omit the details irrelevant here.

¹¹ I turn to this point below soon.

¹² That phonological features themselves do not affect NS operations is claimed in terms of, say reduction of *want to* in 'I want to go to school' to *wanna*: such a reduction concerns only PHON. I turn to the second point, whether presence/absence of phonological features affects NS operations, in the following subsection.

¹³ According to Pesetsky (2000), all *whs* move to [Spec,CP] in NS; it depends on language-particular property whether only one *wh*, all *whs*, some of *whs*, or none of *whs* are spelled out in [Spec,CP]. I turn to *wh*-in-situ later.

¹⁴ Assuming that *wh* passes through [Spec,v*P] (Chomsky 2001), the chain will have one more occurrence of *what*, which I put aside here.

¹⁵ I omit the details here, copies of a subject and an Aux. Assuming a vP for Aux, *what* might pass through [Spec,vP] of Aux, which I omit too.

¹⁶ Concentrating on a chain of *nani-o*, I omit all the other details, the issue of basic word order, a Q-particle, and so forth.

¹⁷ Or alternatively, it could be argued that *wh* in echo question does not move to [Spec,CP], assuming that it does not have [u-wh]. It would be necessary to distinguish *wh* in echo question from the others in that the former does not carry quantificational force while the others do, which I leave aside here.

¹⁸ I turn to head movement later in detail. I counterargue Nunes' account of verb movement.

¹⁹ Nunes seems to assume both [u-wh] and [u-Case]; Chomsky (2000) assumes only [u-wh]. Hereafter, I follow Chomsky for convenience's sake.

²⁰ It could be assumed that the formal features consist of [person], [number], [gender], and so

on., which I put aside here.

²¹ I omit a *v* projection for Aux and a copy of Aux movement.

²² Alternatively, it would be assumed that a pronominal without phonetic realization does not move to [Spec,TP], with *pro*'s formal features deleted in the original position. Another pending issues should be discussed; for instance, validity of assuming an expletive *pro*, or EPP deletion by verb movement (Alexiadou and Anagnostopoulou 1998). I leave detailed analyses for future research.

²³ According to H&M's convention, the substantive lexical forms do not exist at this derivational stage. I notate them as *eat*, and so forth, for convenience' sake. I omit a *v**P projection between VP and T. The subject should leave a copy in [Spec,*v**P], assuming the *v**P projection and the Subject VP-internal Hypothesis (Koopman and Sportiche 1991). As the copy is supposed not to affect syntactic operations (e.g. Chomsky 2000), I put it aside here.

²⁴ In effect, it is not *V* but *v*, a functional head with transitive force, that enters selection relation with T and values an Acc Case (Chomsky 2000~). Here, I notate as *V*.

²⁵ See footnote 9.

²⁶ As for order of functional categories, see Pollock (1989) and Cinque (1999).

²⁷ I omit detailed procedures, assuming the current system (Chomsky 2000~): a probe-goal system, a procedure of Case valuation and deletion, etc. I turn to chain formation later. Here, I focus on the verbal feature complex only. I discuss a chain of verb movement later in detail.

²⁸ I claim later that Agr will enter (a chain) after S-O, while Case should exist from the beginning of a derivation.

²⁹ Assuming that T(+V) moves to be attached to Agr.

³⁰ I assume that *Zayd* and *a house* enter Agree with the matrix T and V, and are valued the Cases. Concerning chain formation and null subject, I turn to in the next section.

³¹ Except an empty subject [*they*] in (38b).

³² As previously, hereafter too I assume mechanisms of the current system: Agree between T and *John*, deletion of T's EPP in terms of *John*'s movement, Case valuation, and so on.

³³ This may be controversial, with wider aspects of information structure of a sentence taken into consideration. It is possible to give a focus accent to a subject in English, as in 'JOHN (, not Bill) did it,' which indicates that the subject can be assigned [Foc] in English. Rizzi (2004) claims, based on the observation of difference among passives, actives, and Clitic Left Dislocation, that passive construction does not seem to be topicalization of an object. His claim will indicate that a feature that is assigned to the higher *John* of the chain (42) may be [Foc]. Either way, what I would like to show here is that a discursal feature like [Top], which will enter after S-O, will function as determining a pronounced position in a chain. I present only one example of possible analyses. I leave detailed surveys of information structure of a sentence for future research.

³⁴ As a case which is not interpreted as a normal *wh*-interrogative, unlike (46a).

³⁵ *Whs* will also have Case, which I omit here.

³⁶ For detailed discussion of focus accent assignment in *wh*-interrogative, see Culicover and Rochemont (1983).

³⁷ Unless the right-most element 'is right-dislocated, in which case it constitutes a distinct intonational phrase' (Zubizarreta 1998:96):

- i) a. ?Qué le compró QUIÉN, a María?
- b. ?Qué bebida trajo QUIÉN, a la fiesta?

³⁸ As done so far, I mark a phrase which receives accent with capital letters of small size, as *QUIÉN*; an accent mark on a certain word in other cases is that used in the Spanish writing system.

³⁹ Denying that a sentence can have multiple foci, Zubizarreta (1998) assumes that focus is carried on by a pair of multiple *wh*-phrases in multiple *wh*-interrogative, and that an accent which falls on one of the *wh*-phrases functions as identifying the pair constituting focus (Zubizarreta 1998:94). I reject her claim, based on Lambrecht (1994). See the following conversation:

i) Who ate what? – John ate an apple,

According to Zubizarreta, *who* and *what* forms a focus pair. However, the answer contains two assertions, *John* and *an apple*, though a proposition cannot express more than one assertion, following Lambrecht. Therefore, I disagree with Zubizarreta in assuming a focus pair of multiple *wh*-phrases.

⁴⁰ Notice that I have referred to actual/phonetic realization only as *accent*, not as *focus accent*. I distinguish focus as a syntactic feature from (focus) accent as a phonetic realization. Thus, a component which is assigned a focus feature may not always be assigned a focus accent too.

⁴¹ It is not clear whether for *wh* to lose focalizing effect means that the *wh* loses quantificational force too. A D-linked *wh*-phrase is associated with topicality (Rizzi:2004); topic is not quantificational, though focus is (Rizzi 1997:292). A reasoning is that a D-linked *wh*-phrase would lose quantificational force. However, a D-linked *wh*-phrase appears to be subject to Weak Crossover, as illustrated below:

i) *Which man_i did you say his_i boss dislikes t_{which man}?

(Lasnik and Stowell 1991:689)

Weak Crossover is assumed to be a diagnostic which distinguishes quantificational A'-binding from non-quantificational one: the former is sensitive to Weak Crossover, while the latter is not (Rizzi 1997:291; see also Lasnik and Stowell 1991). A D-linked *wh*-phrase would then be involved in quantificational A'-binding (though Rizzi assumes that a D-linked *wh*-phrase is attracted by C with both [Q] and [Top] (Rizzi 2004:22,ft.4)). More detailed survey will be required, which I leave for future research.

⁴² One more case to be discussed is a *which*-phrase case like 'which book did which person buy.' Two chains will be formed until S-O, resulting in <which person-~~WH~~, which-person-~~WH~~> and <which book-~~WH~~, which book-~~WH~~>. As for the former, it would be assumed that [-Foc] enters the higher position, which causes phonological features to be assigned to the lower *which person*. Two analyses might be provided for the latter. Assuming [Foc] enters the higher position, the higher *which book* would have phonological features. Based on Rizzi (2004), on the other hand, [Top] could enter the higher position, which would let the higher *which book* have phonological features. Assuming that [Top] enters, a situation would be yielded in which the sentence does not have focus, against Lambrecht's (1994) claim that a sentence must have focus (The same problem might arise for an echo question case like (62b).). More detailed survey will be required for intonation assignment in *which*-interrogative and its information structure, which I leave for future research.

⁴³ Far more detailed survey is necessary for interaction within multiple *wh*-interrogative, interaction between the superiority effects, D-linking, accent assignment to *wh*-phrases, and possibility of either pair-list or single-pair reading, the last of which I do not refer to here. It is standardly claimed that multiple *wh*-interrogative like 'who bought what?' is answered with pair-list reading: 'John bought apples, Mary bought bananas,...'; the Japanese counterpart 'dare-ga nani-o kaimashi-ta-ka?' (who-Nom what-Acc buy-past-Q), on the other hand, is answered with single-pair reading: 'John-ga ringo-o kaimashi-ta' ('John'-Nom 'apple(s)'-Acc 'buy'-past). Hagstrom (1998) raises data of Sinhala (from Sumangala 1992), stating that the Sinhala multiple *wh*-interrogative can only be answered with single-pair reading, with both *wh*-phrases stressed (Hagstrom 1998:67). Hagstrom also raises data of Okinawan (from

Sugahara 1996), mentioning that when the multiple *wh*-interrogative requires pair-list reading, the reading is a D-linked one (though there is minor difference in Q-morpheme attachment between the multiple *wh*-construction with pair-list reading and that with single-pair reading) (Hagstrom 1998:68). I leave detailed survey of the issue for future research.

⁴⁴ Movement of *-ka* would violate the *Head Movement Constraint* HMC (Travis 1984), as Hagstrom admits. He attempts to avoid it, considering movement to be feature attraction: ‘feature attraction drives movement of the closest element with the relevant feature. ... Any head which does not carry this feature is irrelevant’ (Hagstrom 1998:61). I put aside the issue here, simply following Hagstrom in assuming that the HMC does not apply to Q-particle movement.

⁴⁵ I tentatively assume that an adverb is attached to v*P.

⁴⁶ Julien subdivides the projection system more finely, following Cinque. I somewhat simplify the system here.

⁴⁷ I concentrate on head chain only.

⁴⁸ Strictly speaking, Matushansky means with the uniformity of a chain that occurrences of a chain must all have either head status or phrasal status: it is a standard assumption that a head must not move to a Spec position, and vice versa. I have implied with uniformity that each occurrence must possess the same kind of features: even if a phrase, for instance, moves to a Spec, the chain is not uniform if the occurrences have different features.

⁴⁹ I put aside *v* here.

⁵⁰ In my opinion, addition of features does not necessarily mean affixation as language-particular morphological realization, as Matushansky (2004:3) indicates: I suppose that adding a feature does not always mean that a morpheme corresponding to the feature appears in a language. I assume more abstract feature composition, which must cross-linguistically take place on the assumption that NS and SEM are uniform, with difference attributed to PHON.

References

- Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR, word order, V-movement, and EPP-checking. *Natural Language & Linguistic Theory* 16.491-539.
- Aoun, Joseph. 1982. The Formal Nature of Anaphoric Relations. Cambridge, MA: MIT dissertation.
- Belletti, Adriana. 2001. "Inversion" as Focalization. *Subject Inversion in Romance and the Theory of Universal Grammar*, ed. by Aafke C.-J. Hulk and Jean-Yves Pollock, 60-90. Oxford: Oxford University Press.
- Bobaljik, Jonathan David. 2003. A-Chains at the PF-Interface: Copies and 'Covert' Movement. *Natural Language & Linguistic Theory* 20.197-267.
- Bobaljik, Jonathan David and Samuel Brown. 1997. Interarboreal Operations: Head Movement and the Extension Requirement. *Linguistic Inquiry* 28.345-356.
- Borer, Hagit. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris Publications.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. New York: Mouton de Gruyter.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. *Step by Step*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-156. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by Phase. *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*, ed. by Adriana Belletti, 104-131. Oxford: Oxford University Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Culicover, Peter W. and Michael Rochemont. 1983. Stress and focus in English. *Language* 59.123-165.
- Emonds, Joseph. 1978. The verbal complex V'-V in French. *Linguistic Inquiry* 9.49-77.
- Erteschik-Shir, Nomi. 2001. P-syntactic motivation for movement: imperfect alignment in Object Shift. *Working Papers in Scandinavian Syntax* 68, ed. by Christer Platzack, 49-73. University of Lund.
- Fanselow, Gisbert and Anoop Mahajan. 1995. Partial movement and successive cyclicity. *Papers on Wh-Scope Marking. Arbeitspapiere des Sonderforschungsbereichs 340, Sprachtheoretische Grundlagen für die Computer Linguistik, Bericht Nr. 76*, ed. by Uli Lutz and Gereon Müller, 131-161. University of Stuttgart and University of Tübingen.
- Grimshaw, Jane. 1991. Extended projection. Brandeis University, ms.
- Groat, Erich and John O'Neil. 1996. Spell-Out at the LF interface. *Minimal ideas*, ed. by Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and C. Jan-Wouter Zwart, 113-139. Amsterdam: John Benjamins.
- Gussenhoven, Carlos. 1983. Focus, mode, and the nucleus. *Journal of Linguistics* 19.377-417.
- Hagstrom, Paul. 1998. Decomposing questions. Cambridge, MA: MIT dissertation.

- Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. *The View from Building 20*, ed. by Kenneth Hale and Samuel Jay Keiser, 111-176. Cambridge, MA: MIT Press.
- Holmberg, Anders. 1986. Word order and syntactic features in the Scandinavian languages and English. PhD dissertation, University of Stockholm.
- Holmberg, Anders. 1999. Remarks on Holmberg's generalization. *Studia Linguistica* 53.1-39.
- Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.
- Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Cambridge, MA: MIT dissertation.
- Julien, Marit. 2002. *Syntactic Heads and Word Formation*. Oxford: Oxford University Press.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Koopman, Hilda and Dominique Sportiche. 1991. The position of subjects. *Lingua* 85.211-258.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Lasnik, Howard and Tim Stowell. 1991. Weakest Crossover. *Linguistic Inquiry* 22.687-720.
- Matushansky, Ora. 2004. Moving a-Head. *MITWPL* (title, TBA, to appear), ed. by Ken Hiraiwa and J. Sabbagh. Cambridge, MA: MIT Working Papers in Linguistics.
- Miyagawa, Shigeru. 2001. The EPP, Scrambling, and *Wh*-in-situ. *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 293-338. Cambridge, MA: MIT Press.
- Nilsen, Øystein. 2003. *Eliminating Positions: Syntax and Semantics of Sentence Modification*. Utrecht: LOT.
- Nunes, Jairo. 1999. Linearization of chains and phonetic realization of chain links. *Working Minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 217-250. Cambridge, MA: MIT Press.
- Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.
- Ouhalla, Jamal. 1991. *Functional Categories and Parametric Variation*. London: Routledge.
- Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. *The Representation of (In)definiteness*, ed. by Eric Reuland and Alice ter Meulen, 98-129. Cambridge, MA: MIT Press.
- Pesetsky, David. 2000. *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press.
- Pesetsky, David and Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 355-426. Cambridge, MA: MIT Press.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20.365-424.
- Richards, Norvin. 1997. What moves where in which language? Cambridge, MA: MIT dissertation.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris Publications.
- Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. *Elements of Grammar*, ed. by Liliane Haegeman, 281-338. Dordrecht: Kluwer Academic Publishers.
- Rizzi, Luigi. 2004. On the Form of Chains: Criterial Positions and ECP Effects. University of

- Siena, ms.
- Roberts, Ian. 1993. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer.
- Rochemont, Michael. 1978. A Theory of Stylistic Rules in English. PhD dissertation, University of Massachusetts at Amherst.
- Selkirk, Elisabeth O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Sigurðsson, Halldór Ármann. 2003. The Silence Principle. *Grammatik i focus: festskrift till Christer Platzack Vol.II*, ed. by Lars-Olof Delsing, Gunlög Josefsson, Halldór Ármann Sigurðsson, and Cecilia Falk, 325-334. Lund: Institutionen för nordiska språk, University of Lund.
- Sugahara, Mariko. 1996. Shuri-Okinawan *kakari*-musubi and movement. *Formal Approaches to Japanese Linguistics 2 MITWPL 29*, ed. by Masatoshi Koizumi, Masayuki Oishi, and Uli Sauerland. Cambridge, MA: MIT Working Papers in Linguistics.
- Sumangala, Lelwala. 1992. Long distance dependencies in Sinhala: The syntax of focus and WH questions. PhD dissertation, Cornell University.
- Svenonius, Peter. 1994. C-selection as feature-checking. *Studia Linguistica* 48(2).133-155.
- Travis, Lisa. 1984. Parameters and Effects of Word Order Variation. Cambridge, MA: MIT dissertation.
- Uriagereka, Juan. 1999. Multiple Spell-Out. *Working Minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 251-282. Cambridge, MA: MIT Press.
- Watanabe, Akira. 1992. Subjacency and S-Structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1.255-291.
- Zubizarreta, Maria Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.