Stony Brook University

Sentence Adverbs in the Kingdom of Agree

A Dissertation Presented

by

Chih-hsiang Shu

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

August 2011

Stony Brook University

The Graduate School

Chih-hsiang Shu

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Richard K. Larson – Dissertation Advisor
Professor, Department of Linguistics

John E. Drury - Chairperson of Defense
Assistant Professor, Department of Linguistics

Daniel L. Finer
Professor, Department of Linguistics

C.-T. James Huang
Professor, Department of Linguistics
Harvard University

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Sentence Adverbs in the Kingdom of Agree

by

Chih-hsiang Shu

Doctor of Philosophy

in

Linguistics

Stony Brook University

2011

This dissertation offers a novel account of the syntax of sentence adverbs. The need for a new account is clear from the lack of descriptive coverage and theoretical coherence in current work on adverbial syntax. Descriptively, the majority of work has so far neglected the fact that sentence adverbs behave syntactically like typical focusing adverbs. There has been no coherent and let alone comprehensive syntactic analysis of various focus- sensitive adverbs in generative grammar.

The main proposal I make is that sentence adverbs, as well as focusing adverbs in general, are 'inflectional affixes writ large'. In other words, sentence adverbs are derived in the same way as inflectional affixes are derived in syntax. In the current Minimalist framework (Chomsky 2000 *et seq.*), this parallelism implies that both involve the Agree operation. More specifically, I propose that sentence adverbs merge with a verbal or a nominal expression as a result of (i) Match between valued interpretable Mood features (the *probe*) on C^0 and unvalued uninterpretable Mood features (the *goal*) on a lower functional or lexical head, and (ii) Valuation, where the valued interpretable Mood feature assigns a value to the *goal*. In order to realize the Valuation, sentence adverbs merge with the lower head that is the locus of the *goal*, or with the projection of the head, as a result of pied-piping, to some extent similar to the way inflectional affixes are spelled out as affixes in order to realize feature valuation (Chomsky 2001 *et seq*). This

merge operation is 'delayed- Merge', since this kind of merge applies after regular set-Merge that involves the head containing the *goal*.

Support for this analysis comes from three sources. First, there is extensive evidence that sentence adverbs behave like C^0 elements, although their surface syntactic positions are usually lower. This suggests some kind of syntactic dependency between C^0 and lower functional or lexical heads. These preliminary but fundamental facts are discussed in chapter 2. Second, in-depth scrutiny of focus-sensitivity based on the notion of alternatives, and the role focus plays in the syntactic positions of sentence adverbs, provide compelling evidence that sentence adverbs are focus-sensitive adverbs. This property, as discussed in chapter 3, is crucial in determining which constituent enters a syntactic dependency relationship with the C^0. Third, based on the Chomsky's (2001 *et seq.*) current developments of the generative grammar, inflectional affixes are derived by the Agree operation, which include Match, Valuation, and realization of the Valuation. Our treatment of sentence adverbs as 'inflectional affixes writ large' is not only compatible with the theory, but also provides further support for it. These issues are discussed in chapter 4.

The major consequence of this work is to have shown that the theory of sentence adverbs and focusing adverbs is closely connected with the architecture of grammar in general, including the syntax-morphology interface, the syntax-semantics interface, and the Agree operation. There should be much to be gained if we seriously explore the consequences of our findings for the syntax of various other expressions not currently considered to form a natural class with sentence adverbs and focusing adverbs, such as inflectional affixes and clitics.

Table of Contents

List of Abbreviations

BA   Marker of the *ba* construction
CA   Corelative adverb
CAI   Connective/'refutation' mood adverb *cai*
Cl   Classifier
Conj   Conjunction
DE   Pre-nominal modification marker *de*
DAODI   Interrogative 'impatience' mood adverb *daodi*
DOU   Connective/quantitative adverb *dou*
DOUDAI   Interrogative 'impatience' mood adverb *doudai*
Exc   Exclamative mood particle
Exp   Experiential aspect marker
FA   Focusing adverb
FSE   Focus-sensitive expression
FW   'Friendly warning' mood particle
GE   'Play down' mood clitic *ge*
JIU   Connective adverb *jiu*
HAI   Argumentative mood adverb *hai*
KE   Contrastive Mood adverb *ke*
Neg   Negation marker
Pft   Perfect aspect marker
Pfv   Perfective aspect marker
Prt   Mood particle
QIANWAN   Imperative intensifying mood adverb *qianwan*
RF   'Reduced forcefulness' mood particle
SA   'Soliciting agreement' mood particle, sentence adverb
YE   Focusing/connective adverb *ye*
YOU   Existential marker *you*
ZHEN   Intensifying mood/degree adverb *zhen*

Acknowledgements

Many people contributed to this dissertation in various ways. I'd like to thank my advisor and other members of my committee: Richard Larson, John Drury, Dan Finer, and James Huang. Richard is one of my mentors of linguistics as a rational inquiry. From my experience learning from and working with him, I have learned not only how demanding and humbling doing research can and should be, but also how easy and smooth it can be once you identified the source of a problem. A case in point is that he is the one who pointed out to me many important conceptual problems of some of my premises, which I had to tackle before I could set about presenting the main proposals of this thesis. John Drury has worked with me on this project for a relatively shorter time, but his criticisms and suggestions greatly helped me to formulate my theory more coherently, to see what many more possible directions I can go in the future, and to learn to be more ambitious and go for the bigger picture whenever opportunities are present. Dan Finer has always paid attention to the general theoretical perspectives that I may have missed, the extent of explicitness and preciseness of my proposals, and has taught me the importance of the need to think outside the box and to read a wide spectrum of literature on a given subject matter. My outside committee member, James Huang, inspired me to take advantage of my knowledge of Chinese descriptive linguistics so I can 'study the structure of Chinese in the service of generative grammar'. I'd also like to thank John Bailyn, who could not be on my committee due to his sabbatical arrangement, for asking important questions and offering crucial literature when I was working on the QP which led to this thesis and the thesis proposal later, and offering me important questions to think about and consequences to explore after reading the pre-defense draft of this thesis. I have also benefitted from his vast knowledge of issues of A'-dependency, the influence of which can be seen throughout this thesis.

I'd also like to acknowledge that a significant part of this dissertation should be credited to my informants' help with linguistic data that I personally cannot provide. English data mainly come from faculty members and graduate students who are English speakers in the department: Ginny Anderson, Mark Aronoff, Ellen Broselow, Andy Canariato, Sara Catlin, John Drury, Dan Finer, Shawn Gaffney, Ellen Guigelaar, Robert Hoberman, Joy Janzen, Mark Lindsay, Wesley Parker, Poppy Slocum, Marlyn Taylor, and Julie Weisenberg. Poppy Slocum, in particular, has been hounded by me the most by e-mails and therefore offered most of the examples. In addition, I also appreciate Greville Corbett's help with some sentences when he visited SBU. German examples mainly come from Katharina Schuhmann. Some judgments come from Susi Wurmbrand when she visited SBU. Korean examples mainly come from the Korean graduate students in the department: Young-ran An, Jiwon Hwang, Hijo Kang, and Miran Kim. Russian

ix

examples are offered by Andrei Antonenko.

I'd also like to thank people who gave me various useful comments and tips in my preparation of the thesis. I am grateful to Edith Aldridge, who greatly helped me to refine my term-paper writing and to understand the workings of the Agree theory. Heejeong Ko's encouragements and keen observations helped me to rediscover some of my own insights and have confidence to boldly go forward. Francisco Ordónez showed me how one can make sense of complex morphosyntactic data by looking for fine-grained micro-parametric patterns. Among fellow SBU graduate students, Yukiko Asano's comments to one of my presentation's handouts helped my latter presentations to be much more audience-friendly and helped me clarify my thoughts greatly; Carlos de Cuba offered me useful tips on the syntax of adjunct clauses; Yu-an Lu's feedback on my earlier theory of focusing adverbs in Chinese motivated me to find a better alternative that avoids various empirical problems. I also benefitted from discussions with invited speakers to the department: Chris Barker, Greville Corbett, Sabine Iatridou, Phoevos Panagiotidis, David Pesetsky, Mamoru Saito, Akira Watanabe, and Susi Wurmbrand. Feedback from audiences at conferences also contributed to this thesis, particularly those from Guglielmo Cinque at GLOW in Asia VII. Discussions with linguists at various receptions have also helped to improve the thesis, particularly feedbacks and tips from Henry Yungli Chang, Anna Szabolcsi, and Kensuke Takita. I'd also like to thank Yueh-chin Chang and W.-T. Dylan Tsai for inviting me to NTHU, my alma mater, to give a talk on an early version of this thesis. I'm also grateful to Chris Collins, Aritz Irurtzun, and Doris Penka for sending me their manuscripts.

I'd like to express my deep gratitude to all the linguistics faculty members, staffs, and my fellow graduate students at SBU for their friendship, camaraderie, inspirations, and support. I'd also like to thank my first linguistics teachers and mentors, Kuang Mei and W.-T. Dylan Tsai, for sparking my interest in generative grammar and in rational inquiry in general, and for continuing to offer me guidance after I graduated from NTHU. Thanks also go to my NTHU classmates who continue to offer their support and friendship after all the years: Liching Chiu, Honcin Chow, Wei-wen Roger Liao, Chyan-an Arthur Wang, Jiewu Wei, Iris Hsiao-hung Wu, Barry C.-Y. Yang, and Shih-chi Yeh. With regard to making my life in the US easier, thanks also go to James Huang, who has gone out of his way to treat me and some fellow students he knows with a tour of Cornell University campus, to invite us to parties, and to offer me free lodging in his house during an IACL meeting. Thanks to Jane Tang, who kindled my interest in comparative syntax, has offered me hearty encouragements, treated me to a sumptuous banquet dinner, and invited me to social events with leading Chinese linguists.

And finally, thanks to my parents, Kuei-lun Shu and Hui-ching Lee, for their love and support, and for teaching me important life lessons too numerous to list.

1. Introduction

Sentence adverbs[1] are strange linguistic objects and pose substantial problems for syntactic theory. They are called "sentence adverbs" because, intuitively, they modify the whole sentence, or the proposition it expresses. However, their actual syntactic position is not the sentence-peripheral in many cases, as one might expect given their function. Their syntactic positions have been shown to be focus-sensitive, but it is still not clear how and why this is the case. Their semantics properties are not well understood either, and the syntactic literature has in general failed to accommodate this. As typical adverbial adjuncts, their theoretical status is constantly in debate due to conflicting theoretical considerations and little-understood and under-investigated empirical facts of adverbs and adjuncts in general. There are some other facts that are less well-known but equally puzzling, including some morphosyntactic differences among various sentence adverbs within one language and general cross-linguistic syntactic differences. And these are only a few of the puzzles directly related to sentence adverbs. There are many indirect ones, including their interactions with other classes of adverbs, how human languages encode mood in general, etc.

Astrophysicists studying dark matter, dark energy, pulsar stars, etc. believe "it's often the case that the strangest phenomenon ends up teaching us the most about the universe."[2] It is reasonable to make the same assumption for linguistics. Sentence adverbs seem to be one of the strangest phenomena in syntax, and understanding them may indeed have unexpected impact on syntactic theory.

Thus, the main goal of this thesis is this: to investigate the strange properties of sentence adverbs, in the hope of understanding more about human language.

In this chapter I will provide an overview of the 'strange' yet fundamental properties of sentence adverbs, showing why understanding them is necessary and may shed light on current syntactic theory. The scope and structure of the thesis will also be laid out.

---

[1] Also known as sentential adverbs, sentence adverbials, or sentential adverbials in the literature. I will stick to 'sentence adverbs' throughout the thesis.

[2] This is quoted from astrophysicist Alex Filippenko's comment in the show 'The Universe: Strangest Things' on History Channel aired in 2009.

1.1 A problem of inconsistency

The research interest of this thesis mainly stems from an inconsistency in the syntactic theory.

When syntacticians in the generative grammar tradition talk about grammatical morphemes and words that encode discourse-related or clause-typing information such as wh-words, focus-related morphemes, and particles introducing conditionals, the left-periphery at CP level is the typical place to accommodate these elements. For example, the boldfaced expressions in the following sentences are typically analyzed as occupying the C^0 or spec-of-CP, and this is held to be universal in all human languages:

(1)   a. **What did** Mary see?
      b. **Only then did** he feel better.
      c. **If** you can read this, you're too close.

A major motivation behind the universal CP-periphery analysis is the overarching assumption of the generative grammar, neatly summarized by Chomsky (2001):

(2)   *Uniformity Principle*
      In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.

Following (2), and the empirical fact that there are some overt materials preceding the subjects in the types of sentences in (1), a natural conclusion to reach is that the counterparts of (1) in other languages are also CP-peripheral elements, *modulo* 'compelling evidence to the contrary'.

Another major motivation for treating sentences like (1) as universally involving the CP-periphery is a similar kind of principle to (2), often assumed implicitly by generative grammarians.

(3)   *Structural Uniformity Principle*
      In the absence of compelling evidence to the contrary, assume structures to be uniform.[3]

---

[3] Pesetsky and Torrego (2001: 355) has also promoted a similar view: 'Just as investigation of unfamiliar and diverse languages is regularly illuminated by what is already known about other languages, so the investigation of unfamiliar and diverse *structures* within a single language is regularly illuminated by what is already known about other structures within that language. Again and again, one is led to suspect that an apparent peculiarity of some particular structure is just a special case of a phenomenon characteristic of some entirely different structure.'

X-bar theory (Chomsky 1970), *binary branching hypothesis* (Kayne 1984), *Uniformity of Theta Assignment Hypothesis* (Baker 1988), and *Linear Correspondence Axiom* (Kayne 1994), for example, are all motivated by (3): the first argues that phrases of different categories all have the same basic structure; the second argues syntax allows only binary branching; the third one argues elements with identical semantic interpretations occur in identical syntactic positions with regard to the predicates; the fourth one argues all structures are right-descending. Similarly, if we apply (3) to expressions encoding mood and discourse-related information, we can assume that CP is the one and only structure that expresses such information, unless empirical facts tell us otherwise. One such conclusion is actually explicitly stated in Chomsky (2000: 102):[4]

(4)   LIs fall into two main categories, substantive and functional…Take the *core functional categories* CFCs to be C (expressing force/mood), T (tense[5]/event structure), and *v* (the "light verb" head of transitive constructions).

In other words, the category encoding mood/force information that realizes the boldfaced expressions in (1) should be C and nothing else, unless substantive evidence indicates this to be otherwise.

Most current syntactic theories, however, either argue that there are various possibly non-CP positions for mood categories, or they are part of the split-IP system.[6] Ouhalla (1990) and Zanuttini (1991), for example, propose that sentential negation can select either VP or TP as a parametric variation. Ernst (2002: 374) argues that sentence adverbs can either adjoin to TP, T′, or the first functional category below TP. Den Dikken (2006) argues that when *either* serves as a sentential coordinator, it can choose to either attach to the whole sentential conjunct or the contrastive focus embedded in the sentence. Cinque (1999) proposes that functional categories that accommodate sentence adverbs are part of the split-IP system. To the extent that (3) and (4) are on the right track, these departures are undesirable, although understandable. Linguists simply do not have knock-down evidence to show all mood-related categories are base-generated at C, since they don't always, and sometimes, it seems, cannot, occur very 'high'. This is shown by the following examples[7]:

---

[4]   Here LI abbreviates 'lexical item'.
[5]   Chomsky (2008) further argues that tense is an inherent feature of C. T gets this feature by *inherence* from C.
[6]   Exceptions are those exclusively focusing on sentence-initial or sentence-final expressions (e.g. Cheng 1991, Rizzi 1997, 2004, Haegeman 2000a, 2000b).
[7]   Some, but not every English speaker agrees with the judgment of (5e) and (5f). This could be due to dialectal differences. Similar sentences are also discussed in Collins (1988) and Cinque (1999).

(5) a. John ca**n't** play piano.

    b. You eat **either** all of the ice cream, or I punish you.

    c. Mary **obviously** writes well.

    d. He's in **absolutely no** danger.

    e. John likes **probably** most people in this class.

    f. Bill and **probably** Mary can cook well.

    g. Die Studentin hätte das Buch **wahrscheinlich** gelesen   (German)

       the student   has the book probably   read

       'The student has probably read the book.'

    h. Gianni è   ora **forse**   partito   (Italian)

       G.   has now perhaps left

       'Gianni has perhaps left now.'

    i. zhangsan zuotian   zai gongyuan **jingran**   kandao-le lisi   (Chinese)

       Z.   yesterday at park   surprisingly see-Pfv   L.

       'Surprisingly, Zhangsan saw Lisi at the park yesterday.'

(6) a. ?George says that **evidently** Bob has disappeared.[8]

    b. George says that Bob has **evidently** disappeared.

(7) a. *For **apparently** Bob to be sick would worry Harriet.

    b. For Bob to be **apparently** sick would worry Harriet.

(8) a. *Charley was scared by **stupidly** Violet's driving the car off the cliff.

    b. Charley was scared by Violet's **stupidly** driving the car off the cliff.

(9) a. ?I won't come because **probably** my mother is sick.

    b. I won't come because my mother is **probably** sick.

The inconsistency in the syntactic theory mentioned above can be summarized as follows:[9]

(10) Assuming structures to be uniform, there should be rigid correspondence between the semantic function of a linguistic expression and its syntactic structure, so all mood-related expressions should be encoded at C, but in the mainstream theories of a subset of mood-related expressions, including negation, sentential coordinators, and sentence adverbs, it is frequently proposed that these expressions are encoded on some lower categories.

As far as I know, this inconsistency has not been explicitly acknowledged in the literature.[10]

---

[8] Examples (6-9) are from Jackendoff (1972: 66).

[9] It can also be regarded as a syntax-semantics mismatch problem. See more discussions in chapter 2.

[10] The fact that sentence adverbs are more restricted in the clause-initial or pre-subject position distributionally than

This inattentiveness is understandable, if we see it from the perspective of certain dominating assumptions and inherent limitations of recent generative syntactic theories.

First of all, although adjuncts have long resisted satisfactory theoretical account (cf. some discussions in Boeckx 2003 and Hornstein and Nunes 2008), many syntacticians, most notably Kayne (1994) and Cinque (1999), have concentrated on eliminating them as a separate theoretical construct (following the uniformity principle (3)), and therefore have, perhaps unwittingly, ignored the kind of uniformity in (4) that is possibly more relevant.

Second, the notion of syntax-semantics correspondence has not been a major concern of many syntacticians. For them, since syntax is autonomous and not derived by semantic considerations directly, a syntactic theory only needs to focus on syntactic facts. This assumption, however, also does not validate the ignorance of semantic considerations as clues for evaluating syntactic theories. If semantic representations have access to the output of syntactic operations, as is assumed by the generative grammar tradition generally, then there is little reason for syntacticians to ignore semantic facts. It is very likely, in fact, that semantic considerations provide important clues that may drastically reshape our understanding of syntax.

Third, even for those syntacticians whose main concern is not eliminating adjuncts and who are 'semantics-conscious', they simply haven't had adequate tools for a proper analysis of these expressions. This lack of tools involves both syntactic and semantic tools. Syntactically, until very recently, most theories of syntactic dislocation and quantifier scope employed the mechanisms of feature checking, trace theory, and reconstruction. According to the widely-adopted checking theory developed in Chomsky (1995: 289), if there is a formal feature F in a functional category such as C, there are only two possible operations for F to be checked overtly, Merge or Move, as illustrated below:

(11) a. CP   b. CP
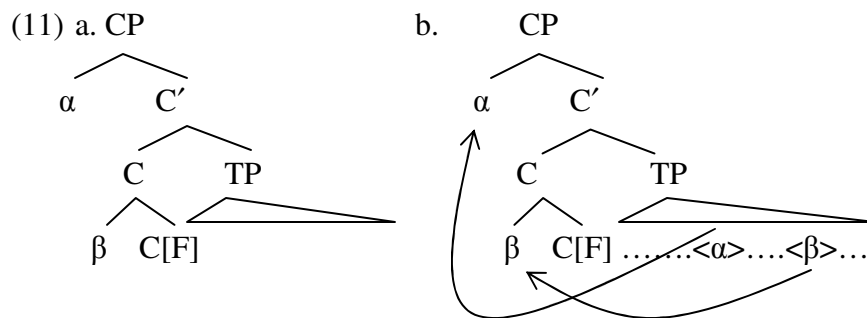
In (11a), operation Merge is involved. Either some maximal projection α is inserted at

spec-of-CP or some β is adjoined to C itself. In (11b), Move is involved. α and β are already present in the derivation before they are chosen to move to these two positions to satisfy the checking requirement. This theory looks simple and straightforward. It also, however, seems to deny any possibilities for negation/sentence adverbs to be involved in checking relations with C. First, as illustrated by the examples in (5), they typically do not need to occur in the sentence-initial position. So an overt Move or Merge analysis that involves CP-periphery in the spirit of (11) is not likely to be an option[11,12]. Second, a covert movement analysis of adverbs has also been generally rejected in the literature[13], due to the fact that they do not manifest scope ambiguities like typical quantified NPs (Ladusaw 1988), and the fact that the base positions for the adverbs are a mystery under the movement analysis, those positions being not θ-related and uncovered by the assumption of the general syntactic architecture (4).[14]

Semantic accounts of sentence adverbs and other mood-related expressions in the literature also leave something to be desired. It hasn't been conclusive among semanticists whether the various mood-related expressions are modifiers, operators, or syncategorematic elements (cf. Jackendoff 1972, Bellert, 1977, McConnell-Ginet 1982). Similarly, it is also noted that these expressions have to be accounted for at least partially in terms of pragmatics (cf. McConnell-Ginet 1982, Ifantidou-Trouki 1993, Potts 2005, Bonami and Godard 2008). Since there have been no concrete proposals of the semantic status of these multifarious expressions, it seems to many syntacticians that there is no need to treat them on a par with those 'force indicators' that syntacticians generally agree occur at C. This solution, however, obviously results from the lack of understanding rather than fruitful research.

Thus, we have reason to believe problem (10) is real. To solve it, one must develop a theory that can not only predict the syntactic distributions of mood-related expressions in (5), but also settle the relevant theoretical problems such as the syntactic status of adjuncts and dislocation of mood-related expressions.

---

[11] One may think it is possible to make this kind analysis work by assuming, rather unconventionally, that all the materials preceding sentential adverbs in (5) are higher than CP. I will discuss and reject this analysis in chapter 2.

[12] Another possibility is treating the 'lower' mood-related expressions as affixes that undergo Affix Hopping in the spirit of Chomsky (1957) and Babaljik (1995), according to which 'I-lowering' is allowed and is a morphological rule applying at PF. Under this analysis, if a mood-related expression at C is an affix, they can 'lower' at PF and attach to the verbal host. This analysis, however, as far as I know, has never been employed to deal with adverbs since they are not regarded as affixes.

[13] On covert movement analyses of English negation and modal auxiliaries, see the Boeckx (2001) (for which covert movement is optional) and Butler (2003) analyses. Their analyses focus on accounting for certain scope facts, but they do not explain why these expressions are merged in lower positions in the first place.

[14] This architecture may be evaluated from a somewhat different perspective in Grimshaw's (1991, 2000) framework. I will abstract away from this perspective.

1.2 Various unsettled issues of sentence adverbs

In addition to issues directly related to problem (10), the study of sentence adverbs by itself is also important, since so many issues are still in need of proper theorizing. These issues include the following: (i) adverbial adjuncts, (ii) unique syntactic distributions, (iii) focus-sensitivity, (iv) heterogeneity, (v) cross-linguistic variation. In what follows I will sketch what these issues are and why they are important.

1.2.1 Adverbial adjuncts

Sentence adverbs have the typical properties of adverbial adjuncts. As a syntactic category, they are traditionally categorized as adverbs, and as tree building material in X-bar style phrase structure, they are categorized as adjuncts. Both terms, 'adverbs' and 'adjuncts', however, are poorly defined and poorly understood.[15]

It is no easy job defining adverbs. Although, as a syntactic category, adverbs are typically defined by their signature *-ly* suffix in English, their co-occurrence with APs, VPs, or AdvPs, and the ability to be coordinated with other adverbial expressions (cf. Delfitto 2006), it is getting more difficult to define them under recent developments of generative grammar, especially under minimalist program. A formal definition has rarely been given in textbooks, and the internal make-up of AdvPs is seldom discussed.[16] This contrasts sharply with nouns and verbs, and their maximal projections NP, VP, DP, which have always figured prominently in the literature. One basic reason for the lack of formal treatment is that adverbs, when they are optional adjuncts and not selected θ-role-bearing arguments, do not select and are not selected, whereas nouns, verbs, prepositions, and even adjectives are.[17] This difference comes from the basic fact that these adverbs are syntactically optional. The problem is that if they do not select and are not selected, then there's no diagnostic to define their syntactic category formally. As for the lack of research on the internal make-up of AdvPs, it is due to a different factor. Adverbs in general do not take complements (Jackendoff 1977)[18], and in many cases do not project at all, especially certain

---

[15] A detailed discussion of relevant issues can also be found in Alexiadou (2002).

[16] As far as I know, only Rubin (1994, 1996, 2003) focus exclusively on the internal make-up of AdjPs, AdvPs and adjunct phrases in general (They are called 'ModPs' in his terminology).

[17] When adjectives are used as predicates, they can select complements, and be selected by copulas or certain verbs. Some examples from English are given below:

  (i)Mary was happy to learn the results.
 (ii)John is certain to win.
(iii)John soon became aware that his efforts were paying dividends.

[18] There are in fact a few exceptions to this generalization, as noted by Pullum and Huddleston (2002: 571):

focusing adverbs which take sentential scope, and some typical sentence adverbs in Chinese. For this reason they are still sometimes termed 'particles'.[19] So far no extensive efforts have been made to give those particles any theoretical status.[20] Therefore, any serious foray into the internal make-up of AdvPs has to wait until this issue is settled.

When it comes to defining adjuncts, one meets no less daunting challenges. The early definition in the X-bar theory looks simple (cf. Chomsky 1986b *et seq*). For example, consider (12):

(12) a. $X'' \rightarrow X''\ Y$
    b. $X' \rightarrow X'\ Y$

In rewrite rules above, Y is the adjunct of X'' and X', respectively. In other words, joining an adjunct to a constituent does not change either the latter's category or bar-level. This simple definition, however, ran into many problems when X-bar theory began to undergo a series of general modifications, sometimes to the point of overhauls. First, Larson's (1988) VP-shell analysis allows the possibility of treating some sentence-final adjuncts in English as the most deeply embedded complements. Second, Pollock's (1989) Split-IP hypothesis, developed by Alexiadou (1997) and Cinque (1999), opens the possibility of treating clausal adjuncts as specifiers of different IPs. Third, in the minimalist framework (Chomsky 1993 et seq.), even rewrite rules such as (12) are no longer adequate. Bar-levels are no longer primitive notions and are derived from their feature composition. If an element is not a projection from smaller constituents, then it is a *minimal* projection. If an element does not project any more, then it is maximal. In other words, all the re-write rules developed in the Government and Binding framework (Chomsky 1981) are reduced to principles related to feature composition and feature

---

 (i) The subsidiary is today operating almost entirely separately from the rest of the company.
 (ii) The duel solves disputes independently of abstract principles of justice.
 (iii) Purchase of State vehicles is handled similarly to all State purchases.
 (iv) Foreign firms in US markets are treated equally with their US counterparts.
 (v) %There were some people who reacted differently than you did.

According to Pullum and Huddleston, there are only 13 such adverbs in English, which is a very small number compare to the number of adjectives that can take complements. It is therefore still fair to say Jackendoff's generalization can be maintained.

[19] Adverbs, prepositions, conjunctions and interjections are all treated as particles, instead of four different parts of speech in Jespersen (1924: 87). Modern syntactic theories have yet to explore implications of this insight.

[20] A few linguists have noted this issue. Bayer (1996: 14), adopting Rothstein's (1991) taxonomy of heads, proposes that focusing 'particles' can be categorized as Type III heads, namely minor functional heads that do not project category features. This move is exactly for the purpose of distinguishing non-projectable particles from projectable adverbs. The particles, however, can project semantic features, so they can trigger operations like negative inversion. It is not clear how this account can be derived from general principles of syntax, and how it fits into theoretical frameworks where semantic features do not participate in syntactic operations.

checking between syntactic objects. Since adjuncts do not participate in feature checking, one therefore no longer has a theory-internal way to define adjunction, and must define it in a construction-specific way (Chomsky 1995: 248, and later Chomsky 2004: 117). In general, new developments in the generative grammar lead or force linguists to develop new theories of adjunction, or *vice versa*, but the new theories so far are still full of problems. Empirical facts that distinguish adjuncts from non-adjuncts noted since pre-minimalist literature (the CED effect, the 'weak island' effect, optionality, recursion, lack of impact on c-selection, counter-cyclicity) pose serious problems for the reductionist approach. The construction-specific approach, on the other hand, has the problem of enriching the theory and still lack of proper accounts of the relevant empirical facts.

For the above reasons, we are sorely in need of a theory of adverbial adjuncts.

1.2.2 Unique syntactic distributions

It has been noted as early as Jackendoff (1972: 49) that English adverbs belong to several distributional classes, sentence adverbs being one of them. Sentence adverbs are distributionally unique in that in general, they only occur in the initial and auxiliary positions[21]. In addition, they must precede other classes of adverbs when the latter also occur in the auxiliary position.[22] This state of affairs is also not accounted for by theories of the distribution of DPs and VPs, which do not have distributional classes. Therefore, adverb-specific theories are proposed to deal with it. The theories proposed so far, however, are beset with various problems in addition to the ones mentioned above.

First, there have been theories (e.g. Travis 1988) proposing that sentence adverbs are 'licensed' by I^0, or C^0, whereas manner adverbs are 'licensed' by V^0. Although this to some extent accounts for their different syntactic distribution, licensing is never explained formally, except that it is different from θ-role assignment. Thus, it seems like restating facts without real explanation.

Second, there are theories (e.g. Cinque 1999) arguing that sentence adverbs undergo feature

---

[21] In English, they can also occur in the sentence-final position if they are preceded by a pause, and in the post-verbal position when the object is a focused QNP. We will abstract away from the former case, but will talk about the latter case in chapter 3.

[22] This is actually not always true. As has been observed in Ernst (2002: 369), sentential adverbs can follow frequency and aspectual adverbs:

 (i) We are still probably north of Princeton.
(ii) And Gretchen Delmere was always certainly an expert on politeness.

At present I abstract away from these cases, since the general pattern still holds. I will return to them in chapter 4.

checking with IP-level functional heads that are not pronounced, whereas VP adverbs are licensed by predication within VP. This also to some extent accounts for the syntactic difference between sentence adverbs and VP adverbs. There are, however, some notable empirical and theoretical issues. First, this kind of feature checking is obviously different from the typical feature checking related to external Merge of arguments and predicates and internal Merge of moved elements, because it neither involves θ-assignment nor movement. Even if adverbial syntax does involve feature checking, overt realization of various relevant heads that undergoes feature checking is never found. It is therefore very difficult to see whether the features are actually checked. These difficulties lead to skepticism about the approach. In theories where long-distance Agree replaces local spec-head agreement, the approach is further cast into doubt.

Third, there are theories that argue the different syntactic distributions of sentence adverbs and VP adverbs come directly from their semantic selectional properties (Ernst 2002). More specifically, adjunction of adverbs is basically free, except that the possible adjunction sites are filtered by semantic factors. VP adverbs are adjoined to VP because the former modify the latter semantically. Sentence adverbs are adjoined to *v*P, TP, T′, or CP because the former select a propositional syntactic object. This analysis may seem to describe the facts, but it is not clear it provides a well-grounded theoretical account. First, this theory runs afoul of the long-standing observation that adverbs do not select, since they are optional and some can attach to almost anywhere. Second, the 'free adjunction' analysis seems too liberal and does not explain why adjuncts differ from non-adjuncts in terms of this freedom.

For these reasons, we need a theory that can properly account for Jackendoff's fundamental observation about the differences between the syntax of sentence adverbs and other classes of adverbs.

1.2.3 Focus-sensitivity

A less well-known but nevertheless real fact about sentence adverbs is that to some extent, they behave like focus-sensitive adverbs, in that their syntactic distributions as well as semantic interpretations can be affected by the position of the focused elements in the sentence. Consider the following example from Krifka (2007):

(13) **Fortunately**, Bill spilled [white]$_F$ wine on the carpet.

According to Krifka, a proper understanding of (13) is as follows:

(14) Among two alternatives, BILL SPILLED RED WINE and BILL SPILLED WHITE WINE, the latter one was more fortunate.

Similar observations can be found in Jackendoff (1972: 252) and König (1991: 12). This property of sentence adverbs is also manifested syntactically. Cinque (1999: 31) notes that most classes of higher AdvPs can be used as focusing adverbs. Both Engels (2005) and Shu (2006) also show that syntactic facts suggest sentence adverbs are focus-sensitive. The general conclusion is that sentence adverbs tend to precede the focused constituent and follow the unfocused materials in languages such as Chinese, English, and German.

This focus-sensitivity property is still not well-understood and so far very few theoretical accounts have been offered, perhaps mainly due to the fact that no satisfactory theory of adjunction is available in the first place, and partially due to the fact that the semantic and syntactic accounts of focus-sensitivity in the literature are still meager[23].

Therefore, we need a theory of adjunction that can cope with focus-sensitivity.

1.2.4 Heterogeneity

Sentence adverbs, even those within the same language, are by no means a homogeneous group with respect to syntactic distribution. The heterogeneity can be observed in two aspects.

First, when two or more sentence adverbs occur in a sentence, they follow fixed word order (Jackendoff 1972, Cinque 1999). For example, Cinque (ibid.: 12) observes this strict ordering of the sentence adverbs: speech act adverbs > evaluative adverbs > modal adverbs. The view of Cinque (1999) is that sentence adverbs are specifiers that merge with different functional categories. The fixed order of these functional categories does not come from results of c-selection, but from a fixed *universal functional hierarchy*. So far, this seems to be the best account available, since it provides a more detailed description of the ordering of sentence adverbs than the other theories, especially the 'free-adjunction' theories. The major problems with this approach come from its implementation, which I have already mentioned above: (i) it is doubtful that sentence adverbs, and adverbs in general, are specifiers; (ii) current theories of Agree allows long-distance agreement, nullifying the idea that adverbs can only be specifiers for them to be licensed; (iii) as we have seen in (5), sentence adverbs can occur in various positions in a sentence, this cannot be easily accounted for by the universal functional hierarchy approach; (iv) the focus-sensitivity property is still unaccounted for. In light of these problems, it seems we

---

[23] There are several explicit syntactic accounts of focus-sensitivity, such as Bayer (1996, 1999), Kayne (1998), Horvath (2007) and Wagner (2009). While these accounts are insightful, they have limited empirical and theoretical coverage and need to be expanded, and they have no analyses of sentence adverbs.

need a different view of the fixed word order among adverbs.

Second, although in most cases, sentence adverbs can occur either in the initial position or the auxiliary position in English, some sentence adverbs can only occur in the auxiliary position. This is already noted in the following examples in Jackendoff (1972: 51):

(15) Albert is **merely/truly/simply** being a fool.
(16) ***Merely/*truly/*simply** Albert is being a fool. (*Truly* ok with a different meaning.)

Similar examples are abundant in Chinese, where many (especially monosyllabic) sentence adverbs cannot occur in the initial position:

(17) a. (***zhen**) zhe-jian chenshan (**zhen**) piaoliang
     ZHEN this-Cl   shirt   ZHEN pretty
     'That shirt is very pretty indeed!'
  b. (***ke**) ni (**ke**) dei dangxin
      KE you KE must beware
     'You really must beware.'
  c. (***cai**) ni (**cai**) shi bendan
     CAI you CAI be fool
     'YOU are a fool. (Contrary to what was assumed in the context)'

These mysterious aspects of heterogeneity must be account for in any good theory of sentence adverbs.

1.2.5 Cross-linguistic variation

Another issue of sentence adverbs is that they show cross-linguistic variations with respect to their syntactic distributions. The variation is usually noted in passing in the literature, and no conclusive theoretical treatment has been provided. Ernst (2002: 374 ff.), who discusses this issue in some detail, has focused on the cross-linguistic variation of adverbs in the 'AuxRange.' His observation is as follows: (i) in English, sentence adverbs can generally occur before or after the first auxiliary; (ii) in French, they occur easily after a second auxiliary; (iii) in Chinese, sentence adverbs occurs before the first auxiliary. His solution to this problem is that the variation comes from how permissive a language is about verb movement. As will become clear later, this approach is based on dubious theoretical assumptions of head movement. Furthermore, there is other syntactic variation among languages that cannot be easily accounted for by verb

movement. This includes some differences between French and Italian, as noted by Belletti (1994: 26), according to whom negation in Italian can adjoin in lower positions than French. The difference between the English data in (5) and the Chinese data in (19) also shows that the head movement account cannot solve everything. Examples in (5e,f), repeated in (18), show that sentence adverbs can occur after the verb and the conjunction in English. In Chinese, these positions are generally not available for sentence adverbs (19).

(18) a. John likes **probably** most people in this class.
     b. Bill and **probably** Mary can cook well.
(19) a. *zhangsan xihuan **dagai**   ban-shang de   meiyi ge tongxue
         Z.   like   probably class-Loc   DE every Cl classmate
     b. *zhangsan he   **dagai**   lisi qu-guo   taipei
         Z.   and   probably L.   go-Exp   Taipei

Plainly, there is more that needs to be said about cross-linguistic syntactic variations of sentence adverbs in any proper theory of sentence adverbs.

It should be clear from the previous discussion that there are many properties of sentence adverbs that still lack theoretical accounts. It is one of the main goals of this thesis to provide a coherent theory that can explain them all.

1.3 Various theories of other phenomena that are related to the study of sentence adverbs

In addition to helping us understand the properties mentioned above, the study of sentence adverbs may very well inform syntactic theories on non-adverbs, and vice versa. Therefore, it is also important that we discuss these theories, and then re-evaluate them, if necessary, when we gain more understanding of sentence adverbs. Some theories that bear on the study of sentence adverbs are those on: (i) V-to-T movement, (ii) weak island effect, (iii) the syntax-morphology interface, and (iv) focus-sensitivity and polarity-item licensing.

1.3.1 V-to-T movement

Pollock's (1989) influential work argues that the Verb-Neg/Adv word order in finite clauses in French and the (Aux-)Neg/Adv-Verb word order in finite clauses in English show that, with respect to main verbs, there is V-to-T movement in French but not in English. This argument is based on the assumption that negation is a head that projects a functional category NegP with a fixed position, and that the relevant adverbs are left-adjoined to VP. This analysis was treated as

a standard textbook analysis, until Chomsky (2001) proposed that head movement be relegated to the phonological component, an idea still under-developed.[24] However, the fact that this analysis crucially relies on the syntax of negation and adverbs (including various sentence adverbs) presupposes that the latter is a settled issue. This is clearly not the case, as the discussion in the previous sections has shown with regard to sentence adverbs.

Objections to some or all analyses of Pollock's verb movement analyses in English and French are centered exactly on this point. Iatridou (1990), citing Di Sciullo and Williams (1987) and Travis (1988), entertains the possibility that adverbs and negation are right-adjoined to the verb, so there is no need to posit V-to-Agr movement. Williams (1994), based on the same idea, further argues that there is also no V-to-T movement in French. The relevant phrase structure is shown in (20):

(20)
```
         VP
        /  \
       V    NP
      / \
     V   Adv/Neg
```
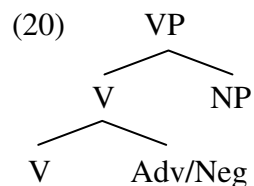
These analyses, however, seem to have intractable problems of their own in the GB era and therefore are not widely accepted. One major problem is their violation of the *structure-preservation hypothesis* that is extended to adjunction. According to this hypothesis, in overt syntax, only YP can adjoin to XP and only Y^0 can adjoin to X^0 (Chomsky 1995: 318). In (20), if the adverb and negation are considered as XPs, which is reasonable since they do not further project, then their adjunction to V violates the structure-preservation principle.

If we have reason to believe, however, that either the structure-preservation hypothesis is no longer valid or that adverbs and negation do not need to be regarded as XPs, then (20) is still a valid analysis for French. These seem to be plausible due to the reasons mentioned in §1.2.1. Furthermore, an analysis like (20) may also help us understand facts mentioned in section §1.2.5 about some less well-known cross-linguistic variations that involve sentence adverbs and auxiliaries.

Therefore, understanding of sentence adverbs should shed important light on the debate as to whether there is V-to-T movement, and would be incomplete without discussions of the latter. On the other hand, theories of V-to-T movement, or theories that argue against it, would also be incomplete without an understanding of sentence adverbs.

---

[24] See Matushansky (2006) for some discussion.

1.3.2 Weak island effect

Another area where theories that involves non-adverbs are relevant is the weak island effect, whose existence has come to be generally recognized since Rizzi (1990). Weak islands refer to a subset of islands that only block extraction of certain phrases. The blockers, or interveners, include wh-islands, certain non-bridge verbs, negation, and certain adverbs. The theoretical question is what exactly makes weak islands weak islands, which is still not properly addressed. It may not be surprising, however, why this is so: weak islands involve configurations with negation and adverbs, including sentence adverbs (Rizzi 2004). Since those elements are still poorly defined theoretically, as mentioned above, it is natural that weak islands are poorly understood.

To see this, we can examine a recent detailed discussion in Rizzi (2004). The article focuses on accounting for the facts within the Relativized Minimality framework and argues that weak island effects can best be derived by RM's version of locality. This locality is defined by the Minimal Configuration (MC). Basically, MC dictates that moving Y to position X is not local if there is an 'intervening' Z of the same structure type as X. Furthermore, due to various types of weak islands that have been discovered so far, Rizzi proposes that there are at least 6 structure types (head, Spec, Argumental, Quantificational, Modifier, Topic). Crucially, he argues that negation and measure adverbs are of the same structure type (Quantificational) as adjunct wh-chains. Manner adverbs belong to a different structure type (Modifier), so they do not block wh-movement. When it comes to adverb preposing, he argues that the 'Modifier' structure type is involved, and all adverbs belong to this class. For this reason, no adverb preposing can cross other adverbs[25].

This approach is rife with problems, although it is one of the most influential, because it is the first work to cover so wide a range of empirical facts. The major problem is the proliferation of 'structure types' is not well-motivated. It has been a central agenda of recent work to derive A and A′ positions from the features involved in these positions (Chomsky 2008: 150), and the results seem to be promising. In Rizzi's analysis, structure types cannot be reduced to features checking or Agree (feature valuation). This simply results in the incompatibility between these two types of theories. Second, even if the theory is right in that structure types cannot be reduced to feature checking or valuation, it is by no means clear what 'structure type' means in Rizzi's framework. The notion is certainly not about the size of the projection (X^0 or XP), the categorical status (CP/AP/DP), or its hierarchical status in a phrase structure. All facts seem to be related to

---

[25] This is not quite true, however. As is noted in Rizzi (1990: 91), temporal, instrumental, and locative adverbs can cross weak islands. He argues that they are in the A position, since semantically they are optionally selected arguments. It is not clear, however, why manner and reason adverbs are not so.

the *semantic* types of the interveners/blockers (which are also not well-motivated). It is far from clear how one can define syntactic notions from purely semantic considerations. These difficulties are directly related to the lack of understanding of the intervening negation and adverbs themselves, and will certainly persist if we stay in the dark about the latter.

1.3.3 The syntax-morphology interface

A theory of syntax-morphology interface is also not complete without an understanding of the syntax of sentence adverbs. It has been noted that some linguistic elements, including certain pronominal elements and adverbs, are 'structurally deficient' in that they are unable to occur in syntactic positions typically allowed for the 'normal' elements (Cardinaletti and Starke 1999). It has also been proposed that certain linguistic elements occur in 'special' positions, separated from the positions reserved for 'normal' elements (Zwicky 1977, Anderson 2005: 31). Most of the studies attempting to address these issues center on pronominal clitics, with syntax-oriented and morphology-oriented linguists working often on diametrically different assumptions. However, very few studies have centered on the morphosyntactic status of the 'structural deficient' or 'special' adverbs. Without those studies, no one can claim that he or she has a theory of 'deficient' or 'special' elements.

Empirical facts about 'structural deficient' adverbs, which are also called 'light' adverbs, in the literature, basically focus on their more restrictive syntactic distributions relative to the verb. Typical examples in English include *simply*, *merely*, *truly* (as we have seen in section 1.2.4), which contrast with other adverbs in their class in that they cannot occur in the initial position. In addition, adverbs such as *not*, *hardly*, *almost*, and *just* contrast with other adverbs in the same semantic class in that the former can only be preverbal (Ernst 2002).

(21) a. The government has (**hardly**) proven its case (***hardly**).
b. The actors might be (**not**) doing their best (***not**).
c. The convoy has (**just**) left (***just**).
(22) a. The lights (**often**) go out (**often**).
b. John (**also**) likes Mary (**also**).
c. (**Now**) he believes the story (**now**).

It is still not clear how syntactic theories can account for the contrast between (15) and (16), (21) and (22), respectively, and (17). The solution provided by the Weight theory in Ernst (2002), which resorts to the feature [+Lite], is not theoretically satisfying since the feature is not well-defined in syntactic theory. Another solution is to give both light adverbs and deficient

pronouns the same morphosyntactic analysis, without resorting to notions of weight, as the one given in Cardinaletti and Starke (1999). This kind of analysis also suffers from theoretical underdevelopment, however, since the theory of 'structural deficiency' is still not yet well-developed or even acknowledged in syntactic theories.

Empirical facts about 'special' adverbs are rarely discussed in the literature. This is not surprising. Since still little is known about the 'standard' positions of adverbs, it is hard to spot 'special' positions of adverbs when we see them. If we can provide an account for the 'standard' positions of adverbs, then we can characterize connections between the standard and the special positions, and thus have a better understanding of the characteristics of clitics.

For these reasons, it is important for the study of sentence adverbs to include morphosyntactic analyses akin to those on pronominal clitics, and also to re-evaluate the analyses of pronominal clitics when more is known about the morphosyntax of sentence adverbs. And more generally, theories of the syntax-morphology interface cannot be complete without studies of both.

1.3.4 Focus-sensitivity

Until recently, theoretical treatments of the syntax of focus-sensitive adverbs are generally ignored or avoided by linguists. This is partially due to the obvious fact that adverbial adjuncts themselves have a murky theoretical status, and partially due to the fact most syntactic theories provide no straightforward ways to accommodate association with focus, or even the syntax of focus in general.

Recently, there has been some progress. First, there has been growing evidence that the syntax of association with *only* and *even* involves covert movement (Krifka (2006), Wagner (2006, 2009)). Second, cases of overt movement of the associates of the focusing adverbs have been discussed to some detail in the context of recent developments of the syntactic theory, including Agree and movement (Horvath 2007). Third, the inventory of focus-sensitive adverbs put under syntactic analyses has grown to include coordinator adverbs such as *either*, *both*, and *neither* (Hendriks (2004), Johannessen (2005), Den Dikken (2006), and Zhang (2008)), and it has been shown these adverbs do have the properties of typical focusing adverbs.

This progress, however, has not yet resulted in an integrated analysis of the syntax of focus-sensitivity. The syntactic processes involved in the insightful analyses of *only* and *even* in Horvath (2007), for example, are not considered in the analyses of coordinator adverbs. More generally, the syntactic analysis of focus-sensitivity is still very limited in terms of theoretical development and empirical coverage, including the possible attaching sites of typical focusing adverbs in languages other than English, the number of focusing adverbs investigated, the reason

why adverbs are attached in the first place, and the relationship between the adverb and the 'associated' elements.

An important task for any linguist studying sentence adverbs is to solve the problems present in these analyses, since they are part of the focus-sensitive adverbs, as has been discussed in §1.2.3. On the other hand, any study of focus-sensitivity is certainly incomplete without acknowledging and studying the focus-sensitivity of sentence adverbs.

To summarize this section, we show that the study of sentence adverbs inevitably involves several syntactic theories of a more general nature. One cannot fully understand one theory without understanding the other. Therefore, the study of sentence adverbs is not only important in its own right, but also important to several areas of syntactic theories that have dominated much of the syntactic literature.

1.4 Road map and scope of the thesis

The main goal of this thesis is to respond to (10), and in doing so address various unsettled issues mentioned in §1.2 and §1.3.

It is necessary to first give a definition of sentence adverbs, for obvious expository and theoretical reasons. This seemingly innocent task will be shown to be more involved than one might think at first, since the term 'sentence adverbs', unlike terms such as subject NPs or embedded CPs, have always been loosely defined and is usually simply assumed in the bulk of the literature. For linguists not using this cover term or using different terms for this class of adverbs, the criteria for their terms are typically based on their assumptions about syntax and semantic in general, not about a universally acknowledged set of syntactic and semantic criteria. Crucially, there have never been any good arguments for or against grouping a set of adverbs into 'sentence adverbs' as a natural class cross-linguistically. Chapter 2 of this thesis is dedicated to this issue, partially as the important first step of setting the scene of the discussions in the rest of the thesis. Theoretical issues concerning what it means to be 'sentence' and what is means to be an 'adverb' will also be addressed. I will conclude that all sentence adverbs should be defined as involving the syntax of C, a definition which is in line with the traditional definitions and modern theoretical frameworks.

Chapter 3 provides the main arguments for treating sentence adverbs as focus-sensitive adverbs, a view that is novel in mainstream syntactic theories, as a second step to understand the syntax of sentence adverbs. The bulk of evidence will come from the restricted positions of sentence adverbs relative to various focused elements in several languages, and also the semantic interpretations of sentences that correspond to different placement of sentence adverbs. Fundamental semantic and syntactic properties of typical focus-sensitive adverbs will also be

reviewed to facilitate our understanding of the syntax of focus-sensitivity. I will establish that the distribution of sentence adverbs, as well as focusing adverbs in general, is regulated by systematic principles that are underrepresented in the literature.

Having laid down basic empirical and theoretical issues of sentence adverbs, I will provide a formal morphosyntactic analysis in chapter 4. I argue for an analysis that involves four processes: Match, Valuation, Pied-piping, and delayed-Merge. Based on the results of chapter 2, I argue that C^0 encodes the interpretable features relevant for the syntax and interpretation of sentence adverbs. Then I will review and sharpen Chomsky's (2000, 2004) theory of Agree, arguing that sentence adverbs are 'inflectional affixes writ large' and are derived just like inflectional affixes are. After the main analyses are given, I go on to investigate several important empirical and theoretical consequences, including cross-linguistic variations of sentence adverbs' attaching sites, cases of overt movement of focused elements, and word order among sentence adverbs, among sentence adverbs and non-sentence adverbs, and among adverbs and non-adverbs. In the end, it will be shown that the main puzzling questions about sentence adverbs that have challenged the linguists so far can find well-motivated solutions from our analyses, while alternative theories on the market have little to say about the relevant facts.

Chapter 5 reviews and explores the consequences of the analysis provided in chapter 4: (i) support for the Agree theory (as opposed to the checking theory), (ii) the NS-Σ mapping is straightforward (there is no syntax-semantics mismatch), (iii) the purpose of Agree is to accommodate the duality of semantics, (iv) an updated trinity of syntax: external Merge, internal Merge, and delayed-Merge, (v) Narrow syntax is not so narrow: support for a fine-grained, sub-modular view of NS. In the outlook section, I will also sketch potential research directions this thesis may lead to.

What this thesis does not and cannot do must also be made clear now. First, it does not provide a comprehensive semantic/pragmatic taxonomy of all sentence adverbs. The task is simply too large[26], and does not seem to have direct bearing on the syntactic analyses presented here. Second, it does not investigate various types of adjunct clauses whose semantic functions seem to be similar to sentence adverbs, but are often realized in the topic positions. The issue is also important, and has been explored to some detail in Haegeman (2003 *et seq.*), but a comprehensive theory that incorporates the results of the present study is beyond the scope of the thesis. Third, it does not investigate various 'parenthetical' adjuncts, whose semantic functions seem to be similar to sentence adverbs, but are marked by a prosodic break. The syntactic status of these expressions has created much debate (see, for example, Haegeman (1991) and Arnold (2007)), but is again beyond the scope of this thesis, since the issues involve syntactic and

---

[26] For a detailed descriptive classification of adverbs in English, see chapter 10 of Biber et al. (1999). For a recent detailed classification of adverbs in Chinese, see Zhang (2000).

phonological considerations orthogonal to the ones discussed here. Fourth, this thesis does not aim to provide accounts of other mood-encoding elements such as verbs, auxiliaries, adjectives, affixes, etc. The latter may be susceptible to morphosyntactic and semantic analyses similar to sentence adverbs, but they are sufficiently different linguistic objects to warrant separate treatment. Finally, the present study only partially explores the syntax of the left-periphery processes that interact with sentence adverbs, including scrambling, topicalization, focus-related movement, various issues of QR, various cases of A-movement, wh-movement, among others, again due to its limited scope.

2. Toward a definition of 'sentence adverb'

In the previous chapter, we saw that sentence adverbs exhibit general properties that continue to lack a satisfactory theoretical account. These include *the inconsistency problem*, the unsettled theoretical status of adverbs and adjuncts, unique distribution of sentence adverbs, their heterogeneous membership, focus-sensitivity, and cross-linguistic variations. Many of these properties have been merely mentioned in the literature but not further explored. There is good reason, however, for not being hasty to address these properties before first putting them in a more general perspective. This is so because 'sentence adverb' as a class of expressions has undergone little rigorous theoretical scrutiny and definition. Furthermore, its unique syntactic status has been challenged in recent years, partially due to lack of progress in the syntactic studies of adverbs and mainly due to an effort by a number of linguists to eliminate the distinction between specifiers and adjuncts. This has lead to the disuse or abandonment of the term by many linguists. If there is no way to give a clear definition to 'sentence adverb', it would seem that any study of sentence adverbs would be a non-starter. For this reason, the focus of this chapter will be on validating the term 'sentence adverb' and providing a clear definition for it.

The task of validating the term sentence adverb will involve purely syntactic considerations, unlike the general practices of semantics and pragmatics-based classifications of adverbs that are prevalent in the literature, which, I believe, are misleading for syntactic studies. Starting from the significance of the attributive noun 'sentence', we show that the set of expressions traditionally regarded as sentence adverbs does show unmistakable syntactic properties of being clausal, much like the properties of expressions typically regarded to be encoded by C^0. These properties will be shown to be effective gatekeepers that allow only sentence adverbs but exclude other classes of adverbs. Next, I attempt to validate the significance of the noun 'adverb'. It will be shown that a number of syntactic facts also unmistakably distinguish sentence adverbs from other clause-level elements, including modal adjectives, modal nouns, and modal verbs, modal affixes. Thus, we derive a set of diagnostics that can clearly identify sentence adverbs. At this

point, a clear definition can then be given for sentence adverbs.

A number of important consequences follow the improved definition of sentence adverbs. We will see that various items that have not been classified as sentence adverbs can now be subsumed in the category. Second, the traditional classification given in Jackendoff (1972) is now updated to accommodate facts not only in English, but more generally. Third, it becomes clear that any theory denying the existence of sentence adverbs will fail to capture the generalizations adduced in this chapter. Fourth, grouping sentence adverbs together predicts that a unified theoretical account for them is possible. This prediction is borne out in the theory proposed in chapter 4.

2.1 Why do we need the term?

Although the term 'sentence adverb' has seen some coverage in the linguistic literature, its current theoretical status in generative grammar is not settled. The component parts of the term, 'sentence' and 'adverb', do not have clear theoretical status, nor does the term itself. If we compare what are relatively well-established to these notions, it is clear to see why. Compared to the term 'sentence', terms like TP and CP have clearly defined syntactic properties such as syntactic distribution, morphological marking, the kind of movements they trigger, etc. Compared to the term 'adverb', which covers the notion 'adverbial adjunct', structural terms like 'specifier', 'head', 'complement', as well as terms of lexical categories such as N, V, P are all well-defined insofar as their roles in syntactic derivations in the current phrase structure theories are clear (specifiers enter the derivation later than the complements, N, V, P have different selectional, case, and agreement properties, etc). The lack of clear definitionf is undesirable since sentence adverbs seem to have some 'real' syntactic properties, just like other kinds of expressions do, but it is still not clear how they should be expressed descriptively and theoretically. Generally speaking, these properties can be subsumed under the following categories:

(1)   General syntactic properties of sentence adverbs
      a. Syntactic properties of adverbial adjuncts.
      b. Syntactic properties associated with higher syntactic positions.
      c. Semantic properties that suggest CP-level syntax.

(1a) covers properties of adverbs and adjuncts, and can be summarized in the following chart:

(2)

Properties of adverbs[1]	Properties of adjuncts[2]
a. Co-occur with APs, VPs, AdvPs, PPs, IPs, CPs, DPs[3].	g. Do not change the category or bar-level of the constituent they are joined to.
b. Often derived from adjectives via a derivational affix (*-ly* in English, *-a* or *-os* in Greek, *-mente* in Spanish, *-weise* in German).	h. Optionality.
	i. Recursion.
	j. Can be left or right adjoined to the target in certain cases.
c. Cannot be stand-alone predicates (license ellipsis, VP-preposing, etc).	k. Occur more distant from the head than complements.
d. Can be coordinated with other adverbial expressions.	l. Can attach at various categorial levels.
	m. Free word order in certain cases.
e. Generally do not select and aren't selected.	n. Apparent counter-cyclicity.
	o. Do not block agreement.
f. Inflection marking is mostly absent.[4]	p. Display the Condition on Extraction Domains (CED) effect.
	q. Display the weak island effects in some cases.

Although sentence adverbs do not have all the properties listed above, it is clear they have the core properties (absence of selection, optionality, apparent counter-cyclicity, etc), which I will discuss in more detail below. They are therefore part of the 'adverbial adjunct' family. These properties have not yet been satisfactorily accounted for in the literature, although their existence is accepted.

(1b) includes the set of syntactic properties that distinguish sentence adverbs from other adverbs and suggest that the former take sentential scope. These properties have not always been

---

[1] For discussion of the properties listed in (2), see Radford (1988), Alexiadou (2002), Boeckx (2003). See also Chang (2001), Chang (2006), Holmer (2006), for cases where adverbs behave like verbs and therefore do not have many of these properties.

[2] I limit my discussions to base-generated adjunction, leaving aside adjunction derived by movement (scrambling, extraposition, right-node raising, etc.), which may or may not be the same phenomenon.

[3] Examples of adverbs attaching to DPs are less well known, but they do exist. Radford (1988: 264) provides examples like *essentially these lines*, *precisely that point*, *nearly all the chocolates*, *rather too many students*, *so few people*, *quite some time*, etc. His definition and analysis of DPs are different from Abney's (1987) analysis, however. I will abstract away from the differences for now.

[4] There are some exceptions. As we will discuss in more detail in chapter 4, there are reasons to believe affixes like *-ly* are agreement markers (cf. also Alexiadou 1997). Other counterexamples include plural marking on adverbs in Korean (Kim 1994), φ-feature agreement in Archi (Kibrik 1994), inflectional-class-marking in Spanish (Harris 1991, Aronoff 1994). In general, however, inflectional marking is more common with other syntactic categories than adverbs.

identified as syntactic in the literature, since surface syntactic evidence is not always clear and linguists have different ideas about whether different classes of adverbs should be distinguished by semantic or syntactic factors. Nevertheless, the state of affairs definitely merits further syntactic investigation, especially in the light of the standard assumption that semantic scope is a function of syntactic position (May 1977, Huang 1982, Reinhart 1984, Kayne 1998, Bruening 2001, *inter alia*). (1b) can be further subdivided as follows:

(3)　a. Higher syntactic positions.
　　　b. Able to scope over the subject of the sentence.
　　　c. Restricted when under the scope of a clausemate sentence modifier/operator.
　　　d. Restricted in embedded clauses.

A well-known observation that concerns (3a) is made by Jackendoff (1972) on sentence adverbs in English. He focuses on three positions in which an adverb can occur: "initial position", "final position without intervening pause", and "auxiliary position (between the subject and the main verb)." According to him, sentence adverbs can be in the initial position and auxiliary position, but not in the final position without pause ((*Certainly*) *John* (*certainly*) *knows Mary* (**certainly*)). VP adverbs, on the other hand, can either occur in all three positions or only in the final position and the auxiliary position ((*Sadly*,) John (*sadly*) dropped his cup of coffee (*sadly*), (**Completely*) *Stanley* (*completely*) *ate his Wheaties* (*completely*)). The distribution facts that separate sentence adverbs from other adverbs are not clear, but it shows that there is a need for the distinction.

Relevant cases that concern (3b) are rarely discussed in the syntactic literature, but this property can be detected when QNP subjects are present. Consider the following examples from McCawley (1988):

(4)　a. 60 percent of the voters probably prefer Dole to Gore.　　　(ambiguous)
　　　b. 60 percent of the voters intentionally left their ballots blank.　(QNO subj. > Adv)
　　　c. 60 percent of the voters completely reject Dole.　　　(QNO subj. > Adv)

(4a) has an interpretation saying 'it is probable that 60 percent of the voters prefer Dole to Gore'. (4b) and (4c), on the other hand, cannot have interpretations where the adverb scopes over the subject. This contrast suggests the syntactic position of *probably* is somehow higher than other adverbs, according to the standard assumption about syntax-semantics mapping.

Examples that are relevant to (3c) have been discussed in some detail in Bellert (1977), which focuses mostly on semantic aspects of adverbs. Some of her examples are in (5)-(7):

(5) a. *Has John probably come?

b. Has John frequently cheated?

(6) a. Fortunately John has come. → John has come.

Fortunately John has not come. → John has not come.

b. John is speaking loudly. → John is speaking.

John is not speaking lóudly. → John is speaking.

(7) a. *Never did John fortunately run so fast.

b. John fortunately never ran so fast.

According to Bellert, the facts in (5-7) suggest that sentence adverbs, unlike manner adverbs and frequency adverbs, take scope over the whole proposition. From a syntax-semantics mapping point of view, this entails that sentence adverbs are higher than phrase markers that express propositions, and higher also than other classes of adverbs.

Facts relevant to (3d), which may be related to (3c), have seen much less discussion in the literature. Taglicht (2001), for example, observes that what he calls 'mild *actually*' cannot freely occur in all embedded clauses.[5] Specifically, it can only occur in those clauses compatible with 'positive assertion'. The examples he gives are quite robust:

(8) a. They think that actually he was informed.

b. *They demand that actually he should be informed.

c. I hope that actually he won the game.

d. *I hope that actually he wins the game.

e. *(He may stay on, but) if actually he leaves, we'll have to replace him.

f. If actually he's leaving us at the end of the week, …

g. His actually being an impostor and not the Harvard graduate we took him for does not alter the fact that he's very good at his job.

h. *Mary was in favour of our actually going by bus (not by train).

Although the facts here are not as straightforward as the ones in (5-7), since there are no obvious overt militant modifiers/modifiers in these embedded clauses, the fundamental issue is the same. Sentence adverbs take wide semantic scope, so they interact with expressions that correspond to propositions, or, more precisely, the mood of the propositions, and therefore they must occupy high syntactic positions with regard to those expressions and the other types of adverbs.

The existence of (1b), therefore, argues for the need for the term 'sentence adverb' as a type

---

[5] Huang and Ochi (2004) have similar observations about the distribution of the interrogative adverbial *the hell*, and Sung (2007) also has similar observations about certain Chinese mood adverbs.

of syntactic expression.

(1c) refers to the semantic properties and the consequent theoretical-internal syntactic motivations for the term. As has been mentioned in chapter one, if we assume structures to be uniform with respect to expressions of a given semantic type (so there is no syntax-semantics mismatch), expressions that express force/mood should consistently be C^0 elements[6,7]. It is well-established since Jackendoff (1972) that sentence adverbs can be generally regarded as speaker-oriented adverbs from the semantic point of view, since sentence adverbs such as epistemic, evaluative, and evidential adverbs explicitly express the speaker's subjective attitude or commitment to the propositional content of a clause. Although the nature of mood still sees little research in modern semantic theories,[8] it is clear that sentence adverbs generally are semantically akin to classical grammatical mood categories such as subjunctive, imperative, and indicative, optative, dubitative, or various mood-related sentence-final particles in Chinese.[9] In fact, sentence adverbs in Chinese have been termed *yuqi fuci* 'mood adverbs' in various recent studies written in Chinese (see Qi 2006 for some discussion). Assuming these adverbs indeed belong to the mood category, we should, in the absence of compelling evidence, treated them as C^0 elements. This therefore constitutes one more reason that we need the term 'sentence adverb' or perhaps the more specific term 'CP-adverb'.

For the above reasons, we still need the long-standing but still poorly-defined term 'sentence adverb' in modern syntactic theories, unless all of the properties in (1) can be shown to be derivable from totally irrelevant, non-syntactic facts (e.g. purely semantic or phonological considerations). With this in mind, we can move on to review how the term has been treated in the literature of syntactic theories.

---

[6] This statement needs to be qualified. In sentences that involve one or more embedded clauses, mood/modality can be expressed by lexical categories such as verbs, predicative adjectives, or nouns. (e.g. *Mary has to go. There is a possibility that Mary will smile. Mary is likely to win. He was reluctant to answer the question.*) The syntax and semantics of these sentences is arguably distinct from monoclausal sentences (cf. Butler 2006, Williams 2009 for some discussions). I will not deal with mood/modal realized as lexical categories in this thesis. As for modal auxiliary verbs, which may not induce a biclausal structure but whose syntactic status is also largely unsettled, I tentatively assume they are T^0 elements but inherit features from C^0 elements, à la Chomsky's (2008) view of tense features.

[7] Rizzi (1997: 284) holds a somewhat different view about how mood is expressed syntactically. For him, mood is a part of the finiteness system that is a core IP-related property but expressed or replicated by the complementizer system. Here I arbitrarily follow Chomsky's (2000) view that mood is an inherent C^0 property (which also includes agreement and tense in Chomsky 2008). Which theory is correct does not affect the need for the term sentence adverb nor the main proposals of this thesis, which is in principle compatible with either.

[8] Cf. Lyons (1995), Quer (2009) for some discussions of this matter.

[9] Jespersen (1924: 321) also treats the adverb *probably* as a mood element for similar considerations.

2.2 Issues related to defining 'sentence adverb' in the syntactic literature

The state of 'sentence adverb' as a theoretically-meaningful term is quite shaky in the literature, despite the justifications reviewed in the previous section. This situation stems from two more general problems: (i) the definition of 'adverbial adjunct' is unsettled; (ii) adverb classification is still unsettled in the theories of adverbs. After recognizing these problems, we can see why 'sentence adverb' is not easily definable.

2.2.1 The definition of 'adverbial adjunct' is unsettled

As mentioned in §1.2.1, the notion 'adverb', when it comes to sentence adverbs, should be understood more precisely in terms of the notion 'adverbial adjunct'. On the one hand, the expressions covered by the term 'adverb' have distinct properties that set them apart from expressions of other syntactic categories. On the other hand, adjuncts also have distinct properties that set them apart from expressions that have effects on the building of phrase structures (such as projecting or changing the bar-level). Those two well-known sets of properties have been summarized in (2). Various approaches in the generative framework have contributed to defining adverbial adjuncts according to the above properties. For easy of exposition and due to the fact that the Minimalist Program and Antisymmetry hypothesis dominate the current generative theoretical climate, I divide these approaches roughly into (i) pre-Minimalist approaches and (ii) AdvP-in-Spec and Minimalist approaches.

2.2.1.1 Pre-Minimalist approaches

The theoretical status of "adjunct" and "adverbial" was relatively straightforward in the literature of generative grammar in the 1980s, as understood in terms of the following X-bar schemata:

(9)   a. $X'' \rightarrow (Y) \, X'' \, (Y)$
      b. $X' \rightarrow (Y) \, X' \, (Y)$

The schemata contain two succinct rewrite rules, where Y is the adjunct of X'' and X', respectively. According to the rules, Y is optional, recursive, can occur on either side of the target, occur more distant from the head than complements, and does not change the category or bar-level of the target. As for the syntactic category Adverb (Adv), its distribution is governed by category-specific versions of (9), such as the following:

(10) a. V″ → (Adv) V″ (Adv)

    b. V′ → (Adv) V′ (Adv)

    c. Adj″ → (Adv) Adj″ (Adv)

    d. Adj′ → (Adv) Adj′ (Adv)

    e. Adv″ → (Adv) Adv″ (Adv)

    f. Adv′ → (Adv) Adv′ (Adv)

     …

While the classic approach described in (9) and (10) can give us correct results of many of the properties described in (2), numerous theoretical and empirical problems ensue. First, it is unclear what regulates the relationships between the syntactic category of the target and the adjunct. (Why does Adv adjoin to V and Adj to N, but not vice versa? Why can P adjoin both to V and to N?) Furthermore, the theory also fails to capture the relationship between the syntactic category of a constituent and its possible positions in an X-bar schema. (Why does Adv typically occur in the adjunct position? Why can Adv typically not occur in the head, complement, or specifier position?) There are also numerous counterexamples to the general properties listed in (2) that cannot be explained by this theory. In many cases adverbs can only attach to the left of their target. (*This phone is very expensive.* vs. **This phone is expensive very.*) Adverbs of the same class cannot occur twice in a sentence. (**Usually John frequently leaves first.*) Adverbs of different classes usually occur with fixed order. (*Bill probably often sees Mary.* vs. **Bill often probably sees Mary.*) Fourth, different types of adverbs are sensitive to verbal, aspectual, or clausal elements, respectively. (**Does John obviously know Mary? *John often knows Mary. *John was immediately sitting in his room.*) These kinds of co-occurrence constraints are also unexpected in the theory.

    Despite these problems, however, and the now-obsolete status of the X-bar schemata, the theory still remains insightful today and cannot be dismissed. Its simple and clear definition for adverbial adjuncts and its capacity to capture many of the basic but diverse empirical facts in (2) have yet to be fully duplicated in modern versions of generative grammar, as we will see shortly.

    Various work based on the Government-Binding (GB) version of the principles-and-parameters approach (cf. Chomsky 1981) have tried to solve individual problems mentioned above by exploring universal lexical properties of adverbs and adjunct-specific principles. According to Travis (1988), for example, adverbs behave differently from expressions of other syntactic categories because of three general principles that regulate their syntax: (i) they are inherently heads that do not project, (ii) they are 'autonomous' theta-markers, and (iii) they are licensed by head features ([AGR] and [Event] on INFL, [Agent] and [Manner] on VERB). These principles apparently cover the empirical facts related to some of the problems mentioned above,

about which simple phrase-structure rules have nothing to say. To illustrate, consider the following examples:

(11) a. proud of their achievements

b. *proudly of their achievements

c. George probably/*completely was ruined by the tornado.

The contrast between (11a) and (11b) is now explained by the universal property of adverbs that they are heads and (typically) do not project. The contrast between two adverbs in (11c) is due to the fact that they are 'licensed' by different features on different heads.[10] It is not clear, however, if the new account is a significant improvement over the old X-bar theory. It requires adverb-specific principles, and introduces a new kind of syntactic licensing not independently motivated. In other words, it's mostly a restatement of empirical facts.

Travis's solution can be considered as a 'supplement approach' to the classical view of adverbial adjuncts, since it does not modify the core assumptions in (9) and (10). Several other approaches, on the other hand, can be considered as 'reductionist approaches' in that they attempt to solve problems by reducing adjunction to other well-behaved grammatical constructs. Besides the more radical approach formulated in the Antisymmetry hypothesis, which will be discussed later, one such approach is offered in Sportiche (1994, 1998). According to this proposal, adjectives and adverbs are in fact not different from the other syntactic categories in that they can project and have complements and specifiers, and there is no such construct as adjunction. In this approach adjectives and adverbs are dominated by a projection whose head takes the modifiee as an argument, that is, either a specifier or a complement, as illustrated in the following examples:

(12) a. John will stupidly answer.

[AdvP [Adv′ [*stupidly*] [VP *answer*]]]

b. John will answer stupidly.

[AdvP [VP *answer*] [Adv′ [Adv *stupidly*]]]

The general intuition behind this approach is that adjectives and adverbs bear the same kind of relation to their modifiee that determiners bear to their NP arguments or predicates to their arguments. The general effect of this approach is that Adv now can occur as a typical head that can take arguments, so it is no longer an anomaly in a phrase structure. It is unclear, however,

---

[10] See §2.2.2.3 for further discussion.

how such approach can naturally accommodate many of the properties in (2) that have been more or less captured by the old X-bar approach, including optionality, free word order, recursion, etc. In addition, this analysis also cannot explain why (12a) can have a subject-oriented reading while (12b) cannot (cf. Jackendoff 1972). The approach thus seems incompatible with various syntactic facts, rather vague about its semantic proposal and is therefore difficult to evaluate.

Apparent counterexamples to the property 'adjuncts occur more distant form the head than complements' lead some linguists to broaden (9) to allow (13).

(13) $X^0 \rightarrow (Y) X^0 (Y)$

This view, which can be regarded as the 'minor adjustment (of the classical theory) approach', can be found in Williams and di Sciullo (1987), Radford, (1988), Sportiche (1988), Iatridou (1990), Williams (1994, 2000), and is also compatible with Travis's (1988) view of adverbial adjuncts. According to this approach, adverbial adjuncts can attach to X^0, in addition to X' and X''. When this happens, $[X^0\ Y]$ becomes a complex word, morphologically and/or syntactically. This analysis is motivated by the following examples from English and French:

(14) a. He isn't proud **enough** of his country.

    b. The weather may turn **out** rather frosty.

(15) a. Jean embrasse **souvent** Marie

       John embraces often   Mary

       'John often embraces Mary.'

    b. Pierre a   vu   **à peine** Marie

       Pierre has seen hardly Mary

       'Pierre has hardly seen Mary.'

    c. Souvent faire **mal** ses   devoirs, . . .

       Often   make badly Poss homework

       'To frequently do one's homework badly'

In theses examples, adverbial adjuncts intervene between the adjectival or verbal heads and their complements. According to this approach, the adverbs are right-adjoined to the verbs/adjectives, as shown in the following structures:

(16)

```
        VP                      AdjP
       /  \                    /    \
     V⁰   NP/AP            Adj⁰     PP
    /  \                  /    \
  V⁰   Adv            Adj⁰    Adv
```
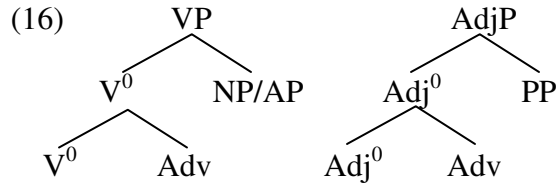
This analysis seems straightforward in so far as it can account for the French (and Italian) facts. However, it has generally been rejected. Pollock (1997: 246), for example, offers two arguments against this view. First, the analysis fails to capture the generalization that if the main verb precedes adverbs in a language, then the verb also precedes negation and undergo V-to-C movement in interrogative sentences, and vice versa. Second, structure (16) also apparently violates the *structure-preserving hypothesis* (SPH) that is extended to adjunction (Chomsky 1995: 318), according to which only YP can adjoin to XP and only Y^0 can adjoin to X^0 in overt syntax. Nevertheless, these two objections are not as strong as they may seem. First, in Shakespearean English, negation can precede the main verb without *do*-support, yet the language has V-to-C movement (van Gelderen 2000, Radford 2004: 150). In Cantonese, the morpheme *dak* 'only' is realized as a postverbal element, yet the language lacks V-to-C movement (Tang 2002). Various Chinese dialects also have the 'potential construction' in which modal elements can be realized postverbally (Chao 1968: 440, Cheng and Sybesma 2003, Huang 2003). None of these dialects have V-to-C movement. On the other hand, the bar-level distinction is also eliminated in the *Bare Phrase Structure* framework (Chomsky 1995), making the SPH vacuous. Counterexamples to the SPH with adjuncts also seem to exist (Toman 1986, 1998, Lieber 1992), such as *an ate too much headache* and *the Charles and Di syndrome*, which appear compatible with an analysis that allows adjoining a YP to an X^0. In addition, the 'main stream' V-to-T or V-to-AGR movement analysis (Pollock 1989), has a serious problem in its central assumption, viz., adverbs occupy fixed positions in the phrase structure. This was shown in chapter one to be inconsistent with the facts of sentence adverbs and other classes of adverbs as well. For these reasons, the adjunction-to-X^0 approach still seems to be a viable addition to the classical approach to adverbial adjuncts, as long as issues of morphosyntax and directionality can be settled.

The fact that some adverbs seem to have multiple attachment sites, often referred to as the 'transportability' property, prompts a line of research that proposes an additional *dimension* to phrase structure. Åfarli (1997), for example, suggests that adverbial adjuncts originate on a z-axis (i.e. beyond the plane) in a 3D phrase structure system, as illustrated below:

(17)

```
              Adv
        XP  /
        / \/
      YP   X′
           / \
          X⁰
```
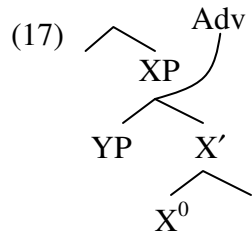
In this approach, the adverbial adjunct still only has one attachment site, the XP, at the initial stage of derivation. However, a later *bending* process allows adverbial adjuncts to linearize with elements 'in the plane.' There are three possibilities for bending, downward bending, leftward bending, and rightward bending, that derives three possible word orders between Adv, YP, and X′, as shown below.

(18) a.

```
      XP
      / \
    YP   X′
         / \
       Adv  X′
```

b.

```
      XP
      / \
    Adv  XP
         / \
        YP  X′
```

c.

```
       XP
       / \
     XP   Adv
     / \
   YP   X′
```
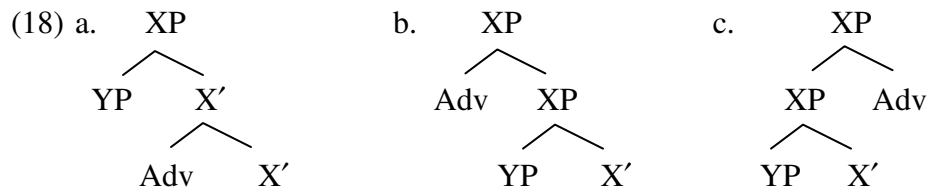
This analysis is motivated partially by Keyser's (1968) observation of English data as follows:

(19) a. John **immediately** sent back the money to the girl.

b. **Immediately**, John sent back the money to the girl.

c. John **immediately** will send back the money to the girl.

d. John will **immediately** send back the money to the girl.

In addition, the analysis is also motivated by facts about apparent two subject and two object positions in Mainland Scandinavian languages, Icelandic, German, and Dutch:

(20) a. Jólasveinninn   borðaði hattin **ekki**   (Icelandic)

the Christmas troll ate   the hat not

b. Jólasveinninn   borðaði **ekki** hattin

the Christmas troll ate   not   the hat

(21) a. Har nogon student **möjligen** last boken?   (Swedish)

has any   student possibly read the book

b. Har **möjligen** nogon student last   boken?

has possibly   any   student read the book

According to Åfarli, the data in (20) and (21) are better analyzed by a transportable-adverb analysis such as the one in (17) instead of a subject-/object-shift analysis. Bobaljik (1999)[11], based on similar facts in Italian and the rigid order of auxiliaries and participles, also argues for a multi-dimensional analysis for adverbs. In spite of the fact that this 'adding dimension approach' can more or less account for the issues of adverb transportability, it also has to be understood as an attempt to preserve certain specific assumptions about adverbs: Åfarli maintains that sentence adverbs only attach to AgrP, and Bobaljik maintains that adverbs occupy fixed positions in the clausal architecture. These assumptions, however, depart from the classical approach's ideas about adverbial adjuncts, according to which the latter can freely attach to VP as well as most other non-nominal categories, which still seem to be valid. Furthermore, the exact workings and consequences of this approach have never been fully explored. It is therefore not clear when and how 3D phrase structure interacts with LF and PF and interface conditions in general, and why 3D phrase structure exists in the first place. All of these problems must be solved before this approach can be pursued further.

2.2.1.2 AdvP-in-Spec approaches and Minimalist approaches

With the advent of the Antisymmetry hypothesis (Kayne 1994), a similar but independently-motivated AdvP-in-Spec analysis (Cinque 1999), and the Minimalist Program (Chomsky 1995), linguists generally found themselves having to reassess radically what they knew about adverbial adjuncts, abandoning a number of old but useful theoretical tools, and redefining adverbial adjuncts in several different ways.

The main proposal of the Antisymmetry hypothesis is the Linear Correspondence Axiom (LCA), according to which linear ordering directly follows from hierarchical relationships (if α asymmetrically c-commands β, then α precedes β). A major consequence of this proposal is that specifiers are taken to be a case of adjoined phrases, and furthermore, each 'specifier' is of a different head. In other words, among the following phrase structures, only (22a) is legitimate.

(22) a. XP   b. XP   c. XP
       / \   / \   / \
     YP   XP   YP   X'   WP   XP
          / \   / \   / \
         X   ZP   X   ZP   YP   XP
                                            / \
                                           X   ZP

---
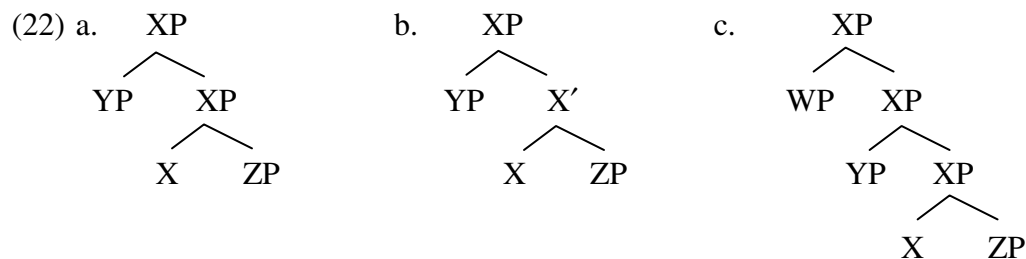
[11] Although Bobaljik's analysis is loosely based on consequences of Cinque's (1999) approach to adverbial adjuncts, he does not exclusively endorse the latter. Therefore, I put it in this subsection.

The major reason for eliminating (22b) is that if it is allowed, the structure will create a problem for LCA and fail to predict the correct word order. This approach to adverbial adjuncts is further developed in Alexiadou (1997) in her analyses of Greek data.

Cinque (1999) also pursues a reductionist analysis, but is motivated by certain specific empirical facts instead of general considerations related to the LCA. Based on his observation that adverbs of different classes cross-linguistically occur in fixed order, and that almost any string of adverbs in a sentence can be interrupted by past participles or finite verbs at any point in Italian, he concludes that there is no adjunction, adverbial phrases are located in the specifier positions of distinct maximal projections. The relevant empirical facts are as follows:

(23) a. Alle due, Gianni non ha **solitamente mica** mangiato, ancora.   (Italian)
      'At two, G. has usually not eaten yet.'
    b. *Alle due, Gianni non ha **mica solitamente** mangiato, ancora.

(24) a. Gianni accetterá **forse saggiamente** il vostro aiuto.
      'G. will perhaps wisely accept your help.'
    b. *Gianni accetterá **saggiamente forse** il vostro aiuto.

(25) a. Gianni ha **per fortuna probabilmente** accettato.
      'G. has luckily probably accepted.'
    b. *Gianni ha **probabilmente per fortuna** accettato.

(26) a. Da allora,   non hanno   **rimesso** di solito mica piú   sempre completamente tutto bene in ordine.
     since then they have put   usually not more always completely all well in order
    b. Da allora, non hanno di solito mica **rimesso** piú sempre completamente tutto bene in ordine.
    c. Da allora, non hanno di solito **rimesso** mica piú sempre completamente tutto bene in ordine.
    d. Da allora, non hanno di solito mica piú **rimesso** sempre completamente tutto bene in ordine.
    e. Da allora, non hanno di solito mica piú sempre **rimesso** completamente tutto bene in ordine.
    f. Da allora, non hanno di solito mica piú sempre completamente **rimesso** tutto bene in ordine.

The examples in (23-25) shows that adverbs of different classes occur in rigid order: habitual adverbs must precede negative adverbs (23), epistemic adverbs precede subject-oriented adverbs (24), and evaluative adverbs precede epistemic adverbs (25). The examples in (26) show that a string of adverbs can be interrupted by an active past participle at any point.[12] Cinque argues that both set of facts, combined with the Pollock's (1989) analysis of French verb-adverb ordering, according to which AdvPs occupy fixed positions and the verb moves, suggest the presence of a distinct head position between the various adverbs in (26), and that the adverbs are

---

[12] Except for the last two, *tutto* and *bien*. Cinque argues that past participles in Italian have to move to the head to the left of *tutto*, but only optionally to the higher heads.

in the specifier position of those heads (à la (22b)). He further argues that the classical adjunction analysis cannot readily account for the facts. One would need an additional 'semantic filter' to derive all the facts, which is redundant in the AdvP-in-Spec analysis, since orders of specifiers simply necessarily follow from a 'universal hierarchy of clausal functional projections'.

The above AdvP-in-Spec analyses focus on certain specific theoretical assumptions and empirical facts about adverbial adjuncts, and appear promising in dealing with issues that are problematic for the classical approach. However, the apparent successes of these reductionist approaches are balanced by significant shortcomings and additional challenges.[13] First, a number of empirical facts that distinguish adjuncts from non-adjuncts that can be directly accommodated by the classical approach no longer follow straightforwardly. It is now difficult to explain why adjuncts in many cases are optional, recursive, can be left- or right-adjoined, and occur further away from the head than complements, etc. If one has to resort more complex and *ad hoc* analyses[14] (e.g. VP-shell analyses, 'roll-up', or 'predicate-raising' analyses, etc.) than the simple classical analyses to deal with these issues, it certainly seems the AdvP-in-Spec approaches are questionable, if not hopeless, as has been suggested in Ernst (2002) and Boeckx (2003). Second, even if we just focus on the facts that are claimed to fare better under the AdvP-in-Spec approaches, such as those in (23-26), we see these approaches still must resort to many *ad hoc* stipulations. For example, the following stipulations are needed to account for (23-26) in the AdvP-in-Spec approaches:

(27) a. Except when V-movement occurs, the heads are phonetically null.
    b. The heads do not block head movement.
    c. Verbal elements can move to those null heads, even though adjunction is barred.
    d. Verbal elements can move optionally.
    e. Adverbs are licensed by null heads.
    f. The specifiers of the relevant null heads need not be realized.
    g. There are a large number of functional projections in each clause.
    h. Feature checking is optional.
    i. Null heads/specifiers are not licensed by any null-element-licensing principles.
    j. AdvP occurs in the spec position of a clausal functional head, but Adv cannot be the clausal head itself.

---

[13] See also Alexiadou (2002: 42 ff) for some general discussions.

[14] This does not mean that VP-shell analyses and predicate raising analyses per se are *ad hoc*, since both may be well-motivated in certain cases. The point here is these analyses do not transparently cover *all* the cases of adverbial adjuncts.

These are stipulations in that none of them are independently motivated in studies of non-adverb specifiers. In fact, none exist in the classical adjunction approach. Thus it seems the AdvP-in-Spec analysis opens the Pandora's Box as far as general issues of null-element-licensing and Spec-head agreement are concerned.[15] Third, there is a general conceptual problem with the "Universal Hierarchy" approach. As noted by Boeckx (2003: 99), the fact that agreement is not subject to the University Hierarchy (noted by Cinque himself and further investigated in Julien 2000) casts doubt on the claim that the University Hierarchy is syntactic in nature, since agreement is an uninterpreted, purely syntactic, property of grammar. If University Hierarchy is semantic in nature, it would be redundant to encode it in the syntax.[16]

In the Minimalist framework, the basic concepts of rewrite rules, bar-levels, categorical labeling, and X-bar theory in general that have been crucial in distinguishing adjuncts from non-adjuncts are replaced by and derived from 'minimalist' concepts and 'virtual conceptual necessity'. These concepts include the following (Chomsky 1995, 2000):

(28) a. The only linguistically significant levels are the interface levels.

    b. The *Inclusiveness Condition*: No new features are introduced by C_{HL} (the computational system).

    c. Relations that enter into C_{HL} either (i) are imposed by legibility conditions or (ii) fall out in some natural way from the computational process.

Bare Phrase Structure (Chomsky 1995) is an explicit theory that aims at deriving the effects of rewrite rules, bar-levels, categorical labeling, and various other GB concepts from these considerations. A consequence of this theory is adjuncts and non-adjuncts cannot be distinguished, and a new kind of labeling is required. That is, the new notations in (29) make the distinction unstatable:

(29) a. X^{max}/XP: A category that does not project any further.

    b. X^{min}: A category that is not a projection at all.

---

[15] Although it seems the same may be said for the classical free adjunction theory, since, as Cinque suggests, a semantic filter is needed to account for the Italian data, the free-adjunction-plus-semantic-filter approach is potentially compatible with Chomsky's (1995, 2000, 2001) analyses of optional operations such as QR and object shift. I will consider details of this alternative in Chapter 4.

[16] Richard Larson (p.c.) also reaches a similar conclusion. He notes that the following parallelisms between monoclausal and multiple-clausal sentences suggest that the rigid ordering is semantic in nature.

(i) a. Jane luckily has probably been granted extra time.
   b. *Jane probably has luckily been granted extra time.
(ii)a. It's lucky for Jane that it's probable that she has been granted extra time.
   b. *It's probable that it's lucky for Jane that she has been granted extra time.

c. Any other category is an X′.

d. Head: A terminal element drawn from the lexicon.

e. Head-complement relation: The most local relation of an XP to a terminal head Y.

f. Head-specifier relation: All other relations within YP (apart from adjunction.)

The problem (29) poses for the adjunct/non-adjunct distinction is that the bar-levels are no longer primitive. Therefore, it does not do to say: 'adjunction of Z to X is a process that gets the label from X but does not change the bar-level of X': in (30), Z can be a specifier, or an adjunct, because bar-level information in not present in the structure.

(30)   X
       /\
      Z   X
         /\
        X   Y
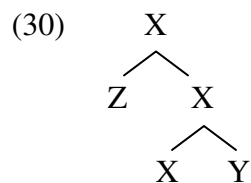
Noting this problem, Chomsky (1995: 248) proposes the following solution:

Substitution forms L={H(K),{α, K}}, where H(K) is the head (= the label) of the projected element K. But adjunction forms a different object. In this case L is a two-segment category, not a new category. Therefore, there must be an object constructed from K but with a label distinct from its head H(K). One minimal choice is the ordered pair ⟨H(K), H(K)⟩. We thus take L={⟨H(K), H(K)⟩, {α, K}}. Note that ⟨H(K), H(K)⟩, the label of L, is not a term of the structure formed. It is not *identical* to the head of K, as before, though it is constructed from it in a trivial way.

This means that adjunction of Z to X now has the following structure:

(31)   ⟨X, X⟩
       /\
      Z   X
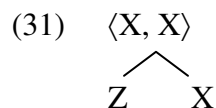
Thus, adjunction is still distinguished from non-adjunction in the early Minimalist Program, now by a new kind of label/category for the result of the adjunction. This approach can be regarded as an 'additional-output-label approach'.

This approach does not deal with the problems for the classical approach noted in §2.2.1.1, but instead focuses on reconciling the Inclusiveness Condition and effects of the classical adjunct/non-adjunct distinction. Inherent problems for the classical approach aside, this

reconciliation does not seem successful, either.[17] First, the label $\langle X, X \rangle$ in (31) is not present in the lexicon, but instead created during the derivation (C_{HL}). Thus the Inclusiveness Condition is still violated. Second, the definition of adjunction in the passage above refers to the notion of "two-segment category". This notion is a pre-Minimalist one and relies on bar-level information, and is therefore incompatible with BPS, making the definition paradoxical.

In later developments of the Minimalist framework (Chomsky 2000, 2004, 2008), a different, more sophisticated minimalist strategy is employed for the treatment of adjunction. Composition of linguistic elements (Merge) is divided into two types, set-Merge and pair-Merge. Set-Merge, dealing with non-adjuncts, forms a symmetrical syntactic object $\{\alpha, \beta\}$ (a set) out of syntactic objects α and β. Pair-Merge, on the other hand, forms an asymmetrical syntactic object $\langle \alpha, \beta \rangle$ (an ordered pair) out of α and β, α being an adjunct (Chomsky 2000). Pair-Merge is also an operation where α is attached to β on a 'separate plane', while β retains all the properties on the 'primary plane' (Chomsky 2004). The essential semantic contribution of pair-Merge is *predicate composition*, not provided by set-Merge. Furthermore, in order to capture the "late-insertion" effects and to permit phonetic linearization, an operation called SIMPL is required, which converts $\langle \alpha, \beta \rangle$ to $\{\alpha, \beta\}$ at the point of Spell-Out (ibid).

According to this approach, example (32) has the derivations shown in (33):

(32) Which picture of Bill$_i$ that John$_j$ liked did he$_{*i/j}$ buy?
(33) a. [$_{NP}$ picture of Bill] (set-Merge forms {picture of Bill})
   b. [$_{NP}$ picture of Bill that John liked] (pair-Merge replaces {picture of Bill} with ⟨that John like, picture of Bill⟩, shadowed elements are on a separate plane)
   c. [$_{CP}$ did he buy [$_{DP}$which [$_{NP}$picture of Bill that John liked]]] (Further set-Merges)
   d. [$_{CP}$ [$_{DP}$Which [$_{NP}$picture of Bill that John liked]]$_i$ did he buy t$_i$] (Wh-movement)
   e. [$_{CP}$ [$_{DP}$Which [$_{NP}$picture of Bill that John liked]]$_i$ did he buy t$_i$] (SIMPL)

In (32), coindexing *Bill* with *he* induces a Binding Condition C effect, but coindexing *John* with *he* does not. A pre-minimalist solution to this contrast is to claim that the NP-adjunct *that John liked* is "inserted late" (Lebeaux 1988). The derivations in (33), however, achieve the same effect without late insertion. First, the NP *picture of Bill* is formed (33a). Next, the NP-adjunct *that John liked* is pair-merged to NP for the purpose of predicate composition, an operation that is cyclic, but does not affect the adjoinee NP and does not create c-command relations between the adjunct and the other elements (33b). Next, the structure [DET $\langle$ADJ, NP$\rangle$] receives its theta role in the normal way, and then further set-Merges occur. There is no Condition C effect when *he*

---

[17] See also Hornstein and Nunes (2008) for similar criticisms of this approach.

and *John* are coindexed because ADJ is not subject to c-command relations at this stage (33c). Wh-movement then applies to produce (33d). Finally, at Spell-Out of the current phase, operation SIMPL removes the "separate plane" and relocates the adjunct to the "primary plane", allowing the adjunct to be phonetically linearized and be subject to further binding relations with materials in the higher phase. Thus, the adjunct is not inserted late, but joins the "primary plane" late.

This approach, which can be viewed as another 'adding dimension' approach, has a number of advantages over the early minimalist approach as well as the pre-minimalist approaches and seems promising, although still with a number of problems. Staying true to minimalist considerations, the effects of adjunct/non-adjunction distinction are derived from the intrinsic properties of lexical items themselves, instead of from extrinsic phrase structure rules and X-bar schemata. This intrinsic property is predicate composition, which is realized by pair-Merge instead of set-Merge. Bar-levels are no longer primitive in this approach. All operations are now cyclic. In addition, the notion of a "separate plane" can capture the effects of transportability as noted by Åfarli (1997), mentioned above, as well as most of the properties of adjuncts in (2). Some problems remain, however. Predicate composition is an ill-understood notion, as noted by Chomsky himself. It is not clear how pair-Merge, and the separate plane, is a consequence of this operation. It is not explained why a separate plane exists and whether it can be derived from something more basic. It is also not clear cross-linguistic differences of adjunction sites can be accounted for in this approach. Typical complex predicate structures have different properties, including their ability to trigger argument structure changes, which are not observed with adjunct structures. In addition, the existence of the optional operation SIMPL seems to add to the *operative complexity* (in the sense of Chomsky 2000: 99), to be avoided unless no better alternatives can be found.

To sum up, I have reviewed various theories of adverbial adjuncts that have attempted to account for their distinctive syntactic properties in various eras of generative grammar, in order to show that its theoretical status remains unsettled. The classical approach provides an X-bar schema that can distinguish between adverbial adjunct and other syntactic expressions pretty well, but the schema itself seems to be derivable from more basic concepts. The approach is also too weak since adjunction is not totally free. The 'supplement' (adverbs are heads licensed by head features) approach makes some progress in empirical coverage, but it relies on a set of stipulations not independently motivated. The 'reductionist' (Adv as a clausal head) approach is a well-meaning attempt at theoretical elegance but falls short of accounting for many of the fundamental properties adverbial adjuncts. The 'minor adjustment' (adjunction-to-X^0) approach looks promising, but invokes theoretical assumptions that lead many syntacticians to favor a verb movement analysis. The early 'addition dimension' approach addresses the transportability

property of adverbial adjuncts, and certain strict word order facts, but it is based on some ill-motivated assumptions and is not clearly spelled out. The AdvP-in-Spec (based on a Universal Hierarchy) approach accounts for the general strict order of adverbs and the Italian word order facts, but is unable to account for most of the properties in (2) and requires many stipulations of its own in order to work. The 'additional-output-label' approach falls short of achieving a real minimalist account. Finally, the 'adding dimension plus SIMPL' approach produces a more specific and coherent analysis of adjunction structures, but leaves many details open and encounters conceptual problems.

2.2.2 Adverb classification is unsettled[18]

We have seen above that the theoretical status of the syntax of adverbial adjuncts is unsettled, despite some promising lines of analysis for a subset of empirical facts. When it comes to syntactic adverb classification, the situation is even worse: relevant theories are much less developed, with conflicting views among linguists.

The underdevelopment of theories of adverb classification is mostly due to inattention. Many studies simply do not care much about adverb classification, and focus instead on theoretical consequences of adverbs on generative grammar as a whole. When attention is shifted to classification, it is usually an afterthought. Second, for studies that do focus on issues of adverb classification, little attention is usually paid to syntactic issues. Studies are more often focused on semantic aspects or isolated syntactic facts of adverb classification, caring more about microscopic perspectives than macroscopic ones. Although these studies sometimes eventually take a stand as to how adverbs should be classified *syntactically*, or at least how they enter syntax in different ways, less vigorous syntactic arguments and analyses are provided than in studies that pay less attention to adverb classification.

The conflicting views of adverb classification generally revolve around whether the classification should be semantic or syntactic in nature. If a theory allows syntactically-free adjunction, then the classification has to be semantic in nature. If a theory argues for syntactic adjunct licensing, then the classification must be at least partially syntactic in nature. Due to these conflicting ontological views, the same diagnostics that distinguish between adverb classes with different distributional properties are said to be "semantic" by some linguists, but are regarded as "syntactic" by others. The situation is aggravated by the fact that, even though

---

[18] Here I limit the discussions to adverbs modifying VP or other clausal categories, which are traditionally referred to as 'adverbials', and abstract away from adverbs solely modifying adjectives, adverbs, PPs, and determiners. However, some 'sentential adverbs in disguise' may attach to the latter types of constituents and will be discussed in the next chapter.

different classes of adverbs can be distinguished by different syntactic and semantic diagnostics, generally all adverbs can occur adjacent to VP or *v*P, at least at surface syntax.

These points will become clearer in the following review of several important works on adverb classification.

2.2.2.1 Jackendoff (1972)

An early influential study addressing adverb classification is Jackendoff (1972). Although developed before the adjunction theory in the 1980s and still use rewrite rules to generate syntactic structures, and hence has little to say about what makes adverbial adjuncts different from non-adjuncts, it does provide a number of detailed observations about the syntax and semantics of different adverbs, how the syntactic and semantic properties are connected, and what a theory of adverbial adjuncts might look like in the generative grammar.

Jackendoff observes that there are three basic positions for adverbs and six distributional classes of them. The three positions are the initial position, the auxiliary position (the position between the subject and the verb), and the final position. The six distributional classes of adverbs are illustrated below:

(34) *Jackendoff's six distributional classes of adverbs in English*
    a. Type I can occupy the initial, auxiliary, and final position, but with meaning changes according to position. (e.g. *cleverly*, *clumsily*, *carefully*, *carelessly…*)
    b. Type II can occupy all three positions, but without discernible change in meaning. (e.g. *quickly*, *slowly*, *reluctantly*, *sadly…*)
    c. Type III can occur in initial and auxiliary position with normal sentence prosody, and in the final position with special prosodic marking. (e.g. *evidently*, *probably*, *unbelievably*, *certainly…*)
    d. Type IV can occur only in auxiliary and final position. (e.g. *completely*, *easily*, *purposefully*, *totally…*)
    e. Type V can occur only in final position. (e.g. *hard*, *more*, *less*, *terribly…*)
    f. Type VI can occur only in auxiliary position. (e.g. *merely*, *truly*, *simply*, *utterly…*)

Jackendoff's proposals about the syntax of various adverbs are as follows. First, as a general theoretical preliminary, he argues that adverb classification is semantic in nature, and syntax just supplies options that semantic rules make reference to. In his system, syntax provides a set of rewrite rules that allow several 'available slots' for adverbs, and semantics provides a set of "projection rules" that make reference to the 'available slots'. As a result, an adverb belongs to a

certain semantic class can only occur in a subset of the 'available slots'. If an adverb occurs in the semantically wrong but syntactically allowed position, then it cannot be interpreted properly by the relevant projection rule, and the sentence will be semantically anomalous. Second, the three basic positions mentioned above correspond to different 'attachment sites' in phrase structure. In initial position, adverbs are attached to S, in auxiliary position they attach either to S or to VP, and in final position they also attach either to S or to VP. Jackendoff also argues that initial position is a derived position, not a base position. Third, based on word order facts, sentence adverbs are analyzed as 'transportable', being able to occur in various positions in a sentence. Fourth, according to some paraphrase possibilities, there are at least three types of adverbs, speaker-oriented adverbs, subject-oriented adverbs, and manner adverbs, the first two have to be "a daughter of S" in order to receive proper interpretation, while the last has to be "dominated by VP" to receive proper interpretation. The relevant rewrite rules and projection rules are illustrated below:

(35) **Rewrite rules**

    a. VP → V NP (Adv*)

    b. VP → (Adv) V NP

    c. S → NP Aux VP (Adv*) (the adverb can be transportable)

(36) **Projection rules**

    Designate the class Adv/PP/S (at least parentheticals)/Modal by F.

    $P_{speaker}$: If F_1 is a daughter of S, embed the reading of S (including any members of F to the right of F_1) as an argument to the reading of F_1.

    $P_{subject}$: If Adv_1 is a daughter of S, embed the reading of S (including any members of F to the right of Adv_1) as one argument to Adv_1, and embed the derived subject of S as tge second argument to Adv_1.

    P_{manner}: Of Adv/PP is dominated by VP, attach its semantic markers to the reading of the verb without changing the functional structure.

This classification is successful for its time, providing detailed descriptions of the syntactic distributions of various adverbs. It also establishes that adverbs have two attachment sites in syntax, VP and S, which nicely fits the well-established semantic properties of different adverbs. However, its attempt to connect syntactic distributions of adverbs and their semantic properties in the generative grammar framework seems to fall short of a coherent solution. A crucial problem for Jackendoff's claim that adverb classification is semantic instead of syntactic in nature is that it is incompatible with Chomsky's (1965) assumptions about *verb* classification,

according to which the subdivision of verbs is expressed in a set of syntactic and semantic lexical features, which Jackendoff also assumes. If syntax plays no role in adverb classification, it is not clear why syntactic features play a role in verb classification, or classification of any syntactic categories.[19] This problem is further exacerbated with later developments of the generative grammar theory, where rewrite rules are eliminated, and syntactic operations rely on syntactic features of lexical items. In addition, Jackendoff's semantics-based classification makes a key misprediction, viz., that adverbs of the same semantic class will have the same syntactic distribution. This prediction is not borne out. Jackendoff himself observes that there are six distributional classes of adverbs. His theory would predict that there are in fact no "distributional classes", but only semantic classes associated with projection rules (the number of which is not six). However, he never reduces the six distributional classes to semantic classes, and it is not clear how to do so. Furthermore, certain sentence adverbs in English, including *truly*, *simply*, *virtually*[20], and many sentence adverbs in Chinese, cannot occur in the sentence-initial position. Jackendoff's transportability theory would predict that they occur freely in the initial position as well as the auxiliary position. Also, it has been shown more recently that sentence adverbs can attach to VP (or PredP) in several languages, such as in French (Ernst 2002: 379), German, and Icelandic (Travis 1988). Jackendoff's theory would predict that these adverbs could only be interpreted as manner adverbs. Finally, as we have already seen in chapter one, sentence adverbs can in fact attach to in various other positions, including object DPs. Jackendoff's theory would predict these are not adverbs, contrary to the fact.

2.2.2.3 Travis (1988)

Travis (1988) develops a more coherent theory in the P&P framework. At this stage of generative grammar, rewrite rules have been replaced by the Principle of Full Interpretation (FI) (Chomsky 1986a). Syntax no longer just assigns 'available slots' to linguistic expressions, but instead requires linguistic expressions be *licensed* by motivated syntactic mechanisms. Travis argues that the consequence of FI for the syntax of adverbs is that a new type of licensing is required: the modifying head (adjective or adverb) is licensed by a feature of the licensing head

---

[19] Another related general problem of this theory is that it has nothing to say about sentences that are syntactically and semantically well-formed but have the wrong interpretations. For example, *John saw Mary* and *Mary saw John* are both well-formed sentences, but they have very different meanings.

[20] These are sentence adverbs by Jackendoff's definition because they are not acceptable in focus-induced subject-aux inversion sentences:

  (i) a. Bill has simply/truly/virtually never seen anything to compare with that.
      b. *Never has Bill simply/truly/virtually seen anything to compare with that.

(noun, verb, etc.). She also refines Jackendoff's observations about 3 positions and 6 distributional classes. According to her, there are four positions allowing adverbs instead of three. These positions include Jackendoff's three and an additional VP-initial position. Travis also makes a slight modification to Type I and Type II adverbs. Type I and Type II adverbs both can be divided into two sub-types, according to their position, since in different positions they have different meanings. The result is that there are now 4 positions and 4 distributional classes (each may be further divided into different semantic classes), as illustrated in the following chart:

(37) <u>Initial/AUX   VP-initial/VP-final   AUX   VP-final</u>

     Type Ia (subject-sensitive)   Type Ib (agent-sensitive)

     Type IIa (event-modifying) Type IIb (process-modifying)

     Type III

                      Type IV

                                        Type V

                            Type VI

Her proposals are as follows. Instead of deriving adverb classes semantically, for example, via projection rules, she argues that they are derived by different licensing heads with different features. Specifically, focusing on the adverbs in the initial/AUX position and those in the VP-initial/VP-final position, she argues that the following features license various adverbs:

(38) V: Agent (Type Ib)

          Manner (Type IIb, IV)

   INFL: AGR (Type Ia)

          Event (Type IIa, III)

   C?:   Speaker

In addition, Travis accounts for variations of syntactic positions within each class (the transportability effect) and cross-linguistic differences by various feature percolation options. According to this theory, the head features in (38) can pass up and down the tree to some extent, subject to parameterization. The process can at most pass down one head, but no more. The INFL features, for example, have the following percolation/transmission options in English, Icelandic, and German:

(39) a. English   b. Icelandic   c. German

```
        IP                    IP                      IP
       /\                    /\                      /\
     NP   I'                NP   I'                 NP   I'
         /\                     /\                      /\
        I   VP               I→VP                   I → VP
                            +F  /\                  +F   /\
                               V   NP                  NP   V
                               +F                           +F
```
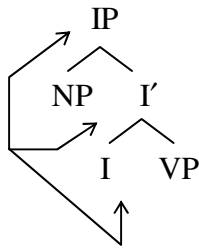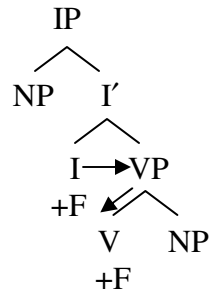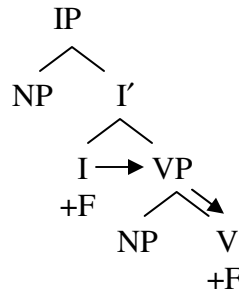
Condition: V movement

In other words, INFL features can percolate upward in English, but cannot pass down to VP. In Icelandic and German, however, the situations is reversed, the features can transmit downward to V, but cannot percolate upward.[21]

This theory of adverb classification represents an improvement over Jackekendoff's approach insofar as the different syntactic behaviors of different adverbs are, like the other syntactic categories, derived by different syntactic features associated with the different classes. It is also able to cover some cross-linguistic facts, for example, where sentence adverbs can attach to VP in Icelandic and German.

However, as mentioned in 2.2.1.1, this theory has various theoretical problems, which become more obvious when we consider the theory's account of adverb classification. First, the notions of head licensing and percolation, while essential for this theory, are in fact only vaguely sketched, and are not independently motivated. If head-licensing is different from predication, as Travis claims, how exactly are they different? What's the nature of the features like [Manner], [Event] of the heads? Are they optional or inherent? What's the semantic effect of head-licensing? It is not clear what adverbs contribute to semantics if it is not a predicate or an argument, or any other known semantic entities. Also, there is no account of percolation. Why does it exist? Does it only apply to head-licensing? If so, why? And how is percolation compatible with FI? How can the cross-linguistic variations be explained? Furthermore, this theory still falls short of making the right factual predictions. It basically has the same predictions as Jackendoff's theory does: adverbs of the same semantic class have the same syntactic distribution, and vice versa, modulo percolation possibilities. These predictions are not borne out. Although there are basically two licensing heads, there are still four classes of adverbs, as we have seen in (37). Although Travis claims Type V adverbs are licensed as prepositions, she doesn't provide evidence for the claim. She also admits for Type VI adverbs she has no analysis. In addition, her

---

[21] Although these facts may be interpreted differently in a CP analysis of verb-second constructions and in various IP-internal functional projection analyses of object-shift effects, the VP-attachment analysis of sentence adverbs is at present still accepted by some linguists (cf. Thráinsson 2000).

percolation theory, although underdeveloped, also makes wrong prediction: as we have seen in chapter one, sentence adverbs can apparently attach to VP[22] and VP-internal constituents in English. This is ruled out in Travis's theory, according to which features involving sentence adverbs cannot be passed down in English, and it cannot be passed down more than one head in any language.

2.2.2.4 Heny (1973)/McCawley (1988)

Another insightful work on adverb classification is McCawley (1988), which expands the observations of Heny (1973). Although its central concern is not the architecture of generative grammar or how adverb classification fits into the picture, it offers some syntactic perspectives that is lacking in other works. McCawley lists four important facts.

First, as has been mentioned in (4), repeated below, some adverbs (40a) can be outside the scope of a quantified NP subject, while other adverbs (40 b,c) cannot.[23]

(40) a. 60 percent of the voters probably prefer Dole to Gore.
    b. 60 percent of the voters intentionally left their ballots blank.
    c. 60 percent of the voters completely reject Dole.

Second, among the classes of adverbs that have to be inside the scope of a QNP subject, one class (41a) can be interpreted as outside the scope of a QNP object, while the other (41b) cannot.

(41) a. Marvin intentionally sliced all three bagels.
    b. The invaders completely destroyed all three villages.

Based on these facts, and the assumption that only S can be the scope of a quantifier, McCawley argues that *possibly* is attached to S, while *intentionally* is attached to a position between S and V′[24], and *completely* to V. Thus, the attachment sites of adverbs can be diagnosed by their surface distributions as well as their scope interactions with QNPs. If we abstract away from McCawley's analysis and match his diagnostics with theories that distinguish between S adverbs and VP adverbs, we have a more or less consistent picture of adverb classification: adverbs that can take scope over QNP subjects are those that attach to S (Jackendoff 1972) or are

---

[22] Here I abstract away from the VP/*v*P distinction, which does not affect the argument here.
[23] Examples involving QNPs basically come from Heny (1973).
[24] His exact formulations of subject-oriented adverbs involve a set of transformational mechanisms of a different theoretical framework than the one assumed in this thesis. I leave them aside here.

licensed by INFL (Travis 1988).[25]

Third, McCawley notes that the class of adverbs called *domain* adverbs (e.g. *linguistically*, *politically*) by Ernst (1984) does not interact with QNPs semantically, but still may occur in different positions with different meanings. This suggests the QNP diagnostic just mentioned is not a necessary condition of the status of sentence adverb. Instead, the higher occurrences of these adverbs have a *domain restriction* function[26].

Fourth, he notes that a set of temporal-related adverbs are also interpreted differently with different positions. These adverbs include *now*, *then*, and *once*[27].

(42) a. John now lives in London. (=in contrast to before)

　　 b. ?? John now is asleep in the next room.

　　 c. John then started shouting at me. (=thereupon)

　　 d. ??I then thought that Philadelphia was in New Jersey.

　　 e. Hemingway once drank at this bar. (=at some time in the past)

He offers no systematic account of the different interpretations of these adverbs, though.

In sum, McCawley (1988) offers a new set of diagnostics on adverb classification and explores several classes of adverbs not explored before in the syntactic literature. Although his focus is not on theoretical development, it certainly enriches linguists' understandings of adverb classes.

2.2.2.5 Cinque (1999)

In the above approaches, adverbs are generally classified into sentence adverbs and VP adverbs, according to the syntactic positions or nodes they are attached to. These approaches are

---

[25] There are some complications here. First, although subject-oriented adverbs in general cannot scope over QNP subjects, some of these adverbs (e.g. *cleverly*, *clumsily*) have similar to other sentence adverbs, as noted by Jackendoff (1972). Second, although all subject-oriented adverbs can scope over the QNP objects, as shown in (41b), only some of them can occur in the sentence-final position with their meanings intact (e.g. *reluctantly*, *intentionally*). (See also Ernst (2002: 54) and Pullum and Huddleston (2002: 676) for some discussions of the two types of subject-oriented adverbs.) These facts show that QNP tests are neither necessary nor sufficient conditions for determining an adverb's syntactic distributions.

[26] Similar observations can be made with adverbs like *slowly* and *quickly*. Although higher occurrences of these adverbs seem to interact with QNPs in sentences like *Slowly, everyone left* and *John slowly sliced three bagels*, this seem to involve ambiguity of a different kind, viz. event-modifying vs. process-modifying, as argued in Travis (1988).

[27] It seems the adverb *long* also has different interpretations in different positions (it preferably occurs in negative sentences).

(i) They haven't long dated each other. (state-modifying?)

(ii) ??They haven't long kissed each other.

based on simple assumptions about phrase structure, where only a few possible attachment sites for adverbs are available (S and VP for Jackendoff and CP, IP, VP, and other 'percolated' positions for Travis, S, V′, and V for McCawley). In late 1980s and early 1990s, however, the situation changed dramatically. As reviewed above, several influential developments (Pollock 1989, Kayne 1994) in phrase structure theory lead to the advent of several AdvP-in-Spec theories, which provided many new 'available slots' for adverbs to attach to in addition to IP and VP. These theories also moved away from a narrow focus on English adverbs, broadening the perspective to include languages such as French, Greek, and Italian. The theoretical and empirical shift of perspective resulted in a new approach to adverb classification.

Cinque (1999) argued that adverbs of different classes are distinguished by different syntactic features, instead of by semantics. Further, Second, Cinque argued that adverbs are not heads, but rather XPs, licensed via spec-head agreement with null functional heads bearing relevant features. The resultant explosion of functional heads and their relative positions include the following:

(43) *The universal hierarchy of clausal functional projections in Cinque (1999: 106)*
[*frankly* Mood$_{speech\ act}$ [*fortunately* Mood$_{evaluative}$ [*allegedly* Mood$_{evidential}$ [*probably* Mod$_{epistemic}$ [*once* T(Past) [*then* T(Future) [*perhaps* Mood$_{irrealis}$ [*necessarily* Mod$_{necessity}$ [*possibly* Mod$_{possibility}$ [*usually* Asp$_{habitual}$ [*again* Asp$_{repetitive(I)}$ [*often* Asp$_{frequentative(I)}$ [*intentionally* Mod$_{volitional}$ [*quickly* Asp$_{celerative(I)}$ [*already* T(Anterior) [*no longer* Asp$_{terminative}$ [*still* Asp$_{continuative}$ [*always* Asp$_{perfect(?)}$ [*just* Asp$_{retrospective}$ [*soon* Asp$_{proximative}$ [*briefly* Asp$_{durative}$ [*characteristically*(?) Asp$_{generic/progressive}$ [*almost* Asp$_{prospective}$ [*completely* Asp$_{SgCompletive(I)}$ [*tutto* Asp$_{PlCompletive}$ [*well* Voice [*fast/early* Asp$_{celerative(II)}$ [*again* Asp$_{repetitive(II)}$ [*often* Asp$_{frequentative(II)}$ [*completely* Asp$_{SgCompletive(II)}$

As for the 'circumstantial' adverbs that follow the verb's complements within the VP, Cinque argues that they are exceptions and he tentatively gives a VP-shell analysis. Third, as a consequence of the existence of fine-grained functional heads sketched in (43) above, no adverb with the same interpretation can occur in more than one position. An adverb occurs only in the specifier position of its licensing head, unless it undergoes topicalization, focus movement, wh-movement, or clitic-climbing-like movement. All other cases of apparently same adverbs occurring in different positions are the results of movements of non-adverbs or XPs containing adverbs, or those adverbs actually having different interpretations. Fourth, the rigid ordering of different classes of adverbs is the result of the universal hierarchy of the licensing functional heads, as shown in (43).

The major motivation, and therefore, strength, of this theory of adverb classification is that

it is able to account for the ordering facts of adverbs not covered in other theories. In addition, it accounts for adverb distribution facts in Italian and French, largely uncharted territory in previous literature. However, there are many problems with regard to adverb classification. First, the syntactic status of (43) is theoretically dubious. As mentioned in 2.2.1.2, there are good reasons to think that the rigid ordering of adverbs could be accounted for by semantics instead of syntax (his universal hierarchy). (43) also to a great extent resembles PS rules in the pre-P&P era, which is not a welcome result in the P&P framework. Second, adverb classification based on ordering facts alone fails to provide a proper account for adverb classification based on other syntactic facts. These include observations in Jackendoff and Travis about several distributional classes of adverbs in English, determined by the position of adverbs with regard to non-adverbs, the observations of McCawley (1988) noted above, and (3b,c). It is not clear how all these follow form the adverb classification in (43), since it makes no predictions about syntactic/semantic co-occurrence restrictions between adverbs and non-adverbs. Third, the 'percolation effect' noted by Travis (1988) is in fact ruled out by this theory, or has to be accounted for by *ad hoc* movement of non-adverbs in German and Icelandic. Fourth, the university hierarchy account itself also provides no account for focusing adverbs and sentence adverbs occurring in low positions in English. Cinque has to resort to free generation of these adverbs as heads taking their modifiees as complements (à la Bayer (1996, 1999)). This account unfortunately undermines the universal hierarchy approach, according to which no such free-generation is possible and all adverbs are XPs.

2.2.2.6 Ernst (2002)

The most recent major study of adverbial syntax and adverb classification is Ernst (2002), which responds to the theoretical problems of the AdvP-in-Spec theories, explores issues of syntax-semantics interface, and aims at accounting for a wider range of empirical facts. To achieve these goals, Ernst reverts to a Jackendoff-style treatment of adverbial syntax, supplemented by an updated version of phrase structure and a somewhat augmented P&P framework.

Ernst's main proposals are as follows. First, in the spirit of Jackendoff (1972), he maintains that syntax itself is quite liberal in providing available adjunction sites, and it is semantic requirements that regulate the distribution of different classes of adverbial adjuncts. Second, he assumes the following classes of adverbs according to the way in which the adjunct combines with its semantic argument (what he calls *Fact-Event Object*, or FEO):

(44) a. predicational

speaker-oriented: Requires a propositional FEO as its argument.

(*frankly*, *maybe*, *luckily*, *obviously*)

subject-oriented: Requires two arguments, an External event and the agent/ experiencer.

(*deliberately*, *stupidly*)

exocomparative: Allows various types of FEOs as its argument. (*similarly*)

event-internal: Requires a Specified Event (a kind of Internal event) as its argument.

(*tightly*, *partially*)

b. domain: Does not take an argument, but imposes a restriction on every predicate.

(*mathematically*, *chemically*)

c. participant: Takes the event variable *e* as its argument.

(*on the wall*, *with a bowl*, *for his aunt*)

d. functional

time-related: location-time, duration, aspectual, (frequency)

Requires an External event as its argument.

(*now*, *for a minute*, *still*)

quantificational: frequency, habitual, additive

Requires an External event as its argument.

(*frequently*, *usually*, *again*, *precisely*)

focusing/clausal-degree: Allows an External event or (indirectly) a proposition as its argument.

(*even*, *just*, *only*, *merely*, *almost*, *nearly*, *just*, *mainly*)

negative: Requires an External event as its argument (*not*, *never*)

clausal relations: purpose, causal, concessive, conditional, etc. (*to win the game*, *if she goes*, *unless they object*, *out of love*, *thus*)

Third, he goes on to offer several syntax-semantics mapping principles and syntax-phonology interface constraints. Some of them are listed as follows:

(45) Constraint on Event-Internal Adverbial Interpretation:

In the domain of L-syntax, only event-internal modification is possible

(46) Extended projection features:

a. [±Disc] = Discourse-related, where [+Disc] heads trigger discourse-related interpretations like topic, focus, and illocutionary force. In the normal case, TP and above are [+Disc].

b. [±C] = Contentful, where only [+C] heads license nonhead items taken from the lexicon, with their own semantic interpretations. TP and below are normally [+C].

(47) T is [+C] for English, [-C] for French and Scandinavian languages.

(48) Checked [+Disc] features on XP add heaviness to XP.

(49) Directionality Principles, including a basic head-initial/head-final parameter.

(50) Weight Theory, which requires, disallows, or (dis)favors certain positions according to weight.

Fourth, he combines the above-mentioned lexicosemantic specifications (44) and general mechanisms such as (45-50) to derive the distribution of different classes of adverbs.

(51) a. predicational

speaker-oriented: Generally occur to the left of nonfinite auxiliaries and negation, modulo (46-48).

subject-oriented: Adjoined to PredP or any higher point that is not higher than T′.

exocomparative: Can occur anywhere.

event-internal: Adjoined to VP or PredP.

b. domain: Can occur anywhere.

c. participant: Adjoined to PredP.

d. functional

time-related: Adjoined to PredP or any higher point, modulo (49).

quantificational: Same as time-related adverbials.

focusing/clausal-degree: Can occur anywhere, modulo (50).

negative: Same as time-related adverbials.

clausal relations: Adjoined to anywhere above VP.

The appeal of Ernst's work lies largely in its detailed exploration of the semantic component of grammar, and its empirical breadth. As can be seen from above, the syntactic distribution of adverbs can be derived without proliferation of functional heads. Semantic composition and certain general syntactic and phonological principles achieve approximately the same results. Empirically, this study broaches issues mostly unaddressed in previous literature, including a more detailed examination of possible attachment sites, issues of information structure, the phonological weight of adjuncts, the directionality parameter; all of these substantiated by a good deal of cross-linguistic data.

There are, however, a number of theoretical and empirical problems that are not easily overcome. Perhaps the most salient theoretical problem is that it still has many of the same problems as Jackendoff's that stems from the attempt to divorce 'syntax' from adverbial syntax. These problems do not simply go away in the modernized framework, as it doesn't follow from

any of the known principles that certain linguistic expressions are licensed by syntactic features while others are not. If adjuncts do not need syntactic licensing since semantics can determine their distributions, it is not clear why non-adjuncts do and why semantics cannot determine the latter's distributions. Second, one of the major motivations of this theory—to reduce adverbial syntax to minimalist interface considerations (p. 3)—seems to stem from a reading of minimalist concepts very different from Chomsky's. In Ernst's focus on dealing with interface considerations, the role of the computational system C_{HL} is largely dismissed (cf. 28). It is simply not true, however, that in the minimalist framework, interface/legibility conditions are 'more important' than computational processes. It is quite reasonable under minimalism that adverbial syntax also heavily involves computational processes. Third, in order to account for facts that cannot be accounted for by purely semantic considerations, Ernst must invoke a set of poorly understood extended projection features such as [±Disc] and [±C]. The involvement of these features makes one wonder how semantics is more important than syntax in adverbial syntax, as Ernst claims, and how this theory is different from Travis's (1988) theory, where adverbs are also licensed by syntactic features of functional heads. Thus, it is not clear if this theory of adverb classification is truly superior to the alternative syntax-oriented approaches.

To sum up, we have seen how adverbs have been classified in the literature, and see that different theories have different strengths and weaknesses. In general, however, the issue is unsettled. Jackendoff (1972) provides a useful descriptive account of six distributional classes of English adverbs, but his semantics-oriented account fails to properly distinguish adverbs from non-adverbs, and makes the wrong predictions about the precise distributions of different classes of adverbs. Travis (1988) provides a more refined version of Jackendoff's distributional classes, and a more coherent theory that derives adverb classification from syntactic considerations, and covers cross-linguistic facts, but the new machineries she proposes — head licensing and percolation — appear *ad hoc* and in need of justification. The theory also continues to make wrong predictions about distributional classes. Cinque (1999) provides detailed account of cross-linguistic adverb ordering facts, but despite its rich theory of functional heads, it still seems unable to deal with old facts such as Jackendoff's observations of six distributional classes of adverbs. In addition, it treats focusing adverbs in a way differently from other adverbs without justification. Ernst (2002) provides a much more detailed account of the semantics and phonology-related facts of different classes of adverbs, but his semantics-oriented theory still faces many of the same problems as Jackendoff's. His ostensibly semantics-oriented analysis appear to be compromised by its heavy appeal to syntactic features such as [±Disc] and [±C], and his account of syntax-semantics interface is based on a very different reading of Chomsky's minimalist framework.

2.2.3 The status of 'sentence adverb' in the literature

From the above review of the literature of adverbial adjuncts and adverb classification, we see that there is no agreement on what an adverbial adjunct is, nor consensus about how adverbs should be classified. Evidently the term 'sentence adverb' has an insecure status in current syntactic theory. Accordingly, when direct reference is made in the literature to sentence adverbs (also referred to as sentential adverbs, sentence adverbials, sentential adverbials), we see two opposing views. According to one view, "sentence adverb" still refers to a real syntactic entity, and is an appropriate object of study. On the other view, sentence adverbs are implicitly or explicitly rejected.

The first view can be seen in various studies of the object shift phenomenon, head movement effects, subject positions, and sentence adverbs themselves. With regard to object shift, sentence adverbs (as well as negation) have been used as a diagnostics for the landing sites at the edge of VP in Thráinsson (2000) and the works cited there. With regard to head movement, sentence adverbs have been used as (i) a marker for the IP-edge position (McCloskey 1996), and (ii) a marker for an Agr$_S$P-edge position (Belletti 1994). With regard to subject positions, sentence adverbs are treated as a diagnostic for clause boundary in bi-clausal structures, which the subject of the lower clause can move across (Postal 1974: 146)[28], and a marker for the T′-edge position in Svenonius (2002). As for studies of sentence adverbs themselves, we see works that focus on them in individual languages such as Jónsson (2002) on Icelandic, Engdahl et al. (2004) on Swedish, and works that touch on some cross-linguistic facts such as Engels (2005) and Shu (2006).

The second view are either implicitly suggested or explicitly expressed in studies that deals with adverb classification discussed above. In Travis's (1988) system, for example, since "S" is no longer a viable syntactic node, she resorts to terms such as INFL-adverbs and COMP-adverbs. In Cinque (1999), sentence adverbs are only mentioned in passing in the context of his much finer-grained adverb classification, and do not have any real theoretical status. In Ernst (2002), for whom there are no different syntactic classes of adverbs, the validity of the term is explicitly challenged (p. 467):

In my view, the frequently invoked terms S(entential) adverb and VP adverb are no longer useful or accurate and, in fact, are quite misleading to the extent that they are meant to express a correlation between adjunction to S/VP and a type of meaning. In some cases, the intended meanings correspond to McConnell-Ginet's Ad-S and Ad-VP

---

[28] Postal's analysis does not directly bear on the syntax of sentence adverbs in monoclausal structures.

or to Jackendoff's speaker-oriented and subject-oriented types, respectively; but in others the intended distinction is between these two as a group ("sentential") and verb modifiers, such as manner or measure adverbs ("VP adverb"). Besides this lack of agreement, the correlation between meaning and adjunction site has never been as close as the terms imply. This was so even before the development of the articulated Infl made up of many functional heads since "VP adverbs" like *shrewdly* may occur before or after subjects (thus being under S (IP)), while "S adverbs" like *perhaps* sometimes show up after one or even two auxiliaries (under VP). With the proliferation of functional heads between subject and V, the inappropriateness of the terms is even more severe.

For these reasons, Ernst resorts to semantic terms such as speaker-oriented adverbs, subject-oriented adverbs, and domain adverbs, etc, despite the fact that those terms are only loosely defined and they share numerous syntactic properties.[29]

To conclude, we have seen that both the theoretical status of adverbial adjunct and a proper delineation of adverb classes are unsettled in the literature, and the term 'sentence adverb' has been either used liberally with little or no justification, or challenged and replaced with other terms despite its intuitive appeal the facts related to (1b,c). In the next section I will propose a solution that can solve these conundrums.

2.3 A modern definition of 'sentence adverb'

In this section, I will show that we can integrate what has been proposed in the literature in a novel way to derive a revealing definition of "sentence adverb." Although details are postponed to the next two chapters, it will be shown that even the preliminary definition is able to provide a more coherent theoretical account as well as wider empirical coverage than previous analyses.

As a first approximation, I propose a definition for sentence adverbs as follows:

---

[29] In more recent work, Ernst (2009) notes that the FEO hierarchy approach is problematic because it fails to predict cases where certain speaker-oriented adverbs can follow negation and question in certain contexts. He thus proposes a new account of adverbial distribution, according to which speaker-oriented adverbs are positive polarity items (PPIs). A surprising consequence, not discussed by Ernst, is that the distinctions between certain speaker-oriented adverbs, subject-oriented adverbs, domain adverbs, etc. are lost, since all of them manifest properties of PPIs, as has been noted in Bellert (1977) and others. More generally, it is not clear how exactly a PPI-based classification can replace or augment a scope-based classification. I will argue below that these properties are partially syntactic properties and can be accounted for in a syntax-based classification.

(52) A linguistic expression α is a sentence adverb if and only if

    a. α has properties of adverbial adjuncts;

    b. α has properties of C^0 elements.

There is nothing new in the definition (52) if (52a) and (52b) are viewed separately. An expression can have properties of adverbial adjuncts such as those listed in (2), which distinguish it from non-adverbial adjuncts or non-adjuncts. An expression can also have properties that are characteristic of C^0 elements, so that it is different from T^0, V^0, D^0, or various non-adverbial XPs. What is new is that (52) says there are expressions that have *both* sets of the properties. In what follows I will show how sentence adverbs fit this new definition. The bulk of data will come from English, with some additions from Chinese.

2.3.1 Sentence adverbs have properties of adverbial adjuncts

    Although sentence adverbs do not have all of the properties listed in (2) (repeated below in (53)), it is clear that they have the core properties. Let's now examine the properties in (2) that typical sentence adverbs have, which are marked with a checkmark. Those not clearly manifested on sentence adverbs are marked with a question mark.

(53) *Properties of adverbial adjuncts shared by sentence adverbs*

Properties of adverbs	Properties of adjuncts
a. ✓Co-occur with APs, VPs, AdvPs, PPs, IPs, CPs, DPs.	g. ✓Do not change the category or bar-level of the constituent they are joined to.
b. (✓)Often derived from adjectives via a derivational affix (*-ly* in English, *-a* or *-os* in Greek, *-mente* in Spanish, *-weise* in German).	h. (✓)Optionality.
	i. ?Recursion.
	j. ?Can be left or right adjoined to the target in certain cases.
c. ✓Cannot be stand-alone predicates (license ellipsis, VP-preposing, etc.).	k. (✓)Occur more distant from the head than complements.
d. ?Can be coordinated with other adverbial expressions.	l. (✓)Can attach at different categorial levels.
e. ✓Generally do not select and aren't selected.	m. ?Free word order in certain cases.
f. ✓Inflection marking is mostly absent.	n. ✓Apparent counter-cyclicity.
	o. ✓Do not block agreement.

| | p. Display the Condition on Extraction Domains (CED) effect. |
| | q. Display the weak island effects in some cases. |

2.3.1.1 Co-occur with APs, VPs, AdvPs, PPs, IPs, CPs, DPs

(53a) is attested for many English speakers (although there are variations of judgment in some cases), and in a more limited way in Chinese, as shown in the following sentences:

(54) a. Mary is rich and **probably** proud of it. (AdjP)

    b. He will **certainly** talk to Mary and **probably** ignore Peter. (VP)

    c. **Fortunately** about the wine, and **regrettably** about everything else, you are correct. (PP)

    d. **Luckily**, he knows everything. (TP)

    e. John likes **probably** most people in this class. (DP)

    f. We can talk about business, **particually** if it will help.[30] (CP)

(55) a. zhangsan yiqian   **yiding**   hen jiaoao (AP)

      Z.   previously certainly very proud

      'Zhangsan previously must be very proud.'

    b. ta **yexu**   renshi   lisi (VP)

      he   perhaps   know   L.

      'Perhaps he knows Lisi.'

    c. ta **xianran**   zai jia (PP)

      he obviously   at   home

      'He's obviously at home.'

    d. **xianran**,   ta zai jia (TP)

      obviously he at home

      'Obviously, he's at home.'

The facts in (54) and (55) may be analyzed in different ways according to assumptions of different syntactic theories. They are plainly compatible with an analysis in which sentence adverbs can attach to various syntactic categories, which will be presented in chapter 4.

---

[30] Although *particularly* is typically classified as non-CP-level degree adverbs, the fact that in (54f) it modifies a conditional clause grants it the status of a sentence adverb.

2.3.1.2 Often be derived from adjectives via a derivational affix

(53b) is attested to some extent in English, and many Indo-European languages. However, it is not generally attested in Chinese. This is shown in the following examples:

(56)   **English**   **Chinese**

Adjective	**Adverb**	**Adjective**	**Adverb**	
surprising	surprising*ly*	yixiangbudao	jingran	
obvious	obvious*ly*	mingxian	xianran	
probable	probab*ly*	dagai	dagai	
fortunate	fortunat*ely*	xingyun	xingyun*de*	
---------	therefore	-----------	yinci	
---------	indeed	-----------	guoran	
---------	perhaps	-----------	yexu	
---------	after all	-----------	bijing	
---------	nevertheless	-----------	yiran	
---------	---------	-----------	suiran	'although'
---------	---------	-----------	nandao	'Can it be…?'
---------	---------	-----------	haishi	'still/had better'
---------	---------	-----------	jiushi	*'even if/absolutely'*

In the above examples, we see that although some sentence adverbs in English are derivationally related to adjectives, most sentence adverbs in Chinese are not morphologically related to an independent adjective form.[31] Thus, (53b) is not a reliable diagnostic for determining whether an expression is a sentence adverb.[32]

2.3.1.3 Cannot be stand-alone predicates (license ellipsis, VP-preposing, etc.)

(53c) is also generally attested. Sentence adverbs do not behave like verbal functional heads

---

[31] There are a number of adverbial suffixes that are not productive in modern Chinese, and should perhaps be analyzed as an inherent part of the adverb, including -*ran* seen in (56). See Tang (1992: 26) for some discussions.
[32] German -*weise*, Greek -*os* are sometimes associate with sentential readings of adverbs (Alexiadou 2002). Japanese -*mo* (See Tamori 1979, Bisang 1998) and Korean -*to* (Heejeong Ko, Jiwon Hwang, p.c.), which are generally regarded as focus particles, also have a similar function. (e.g. *kounni-mo* 'fortunately', *igai-ni-mo* 'unexpectedly', *tahaynghi-to* 'fortunately', *ama-to* 'probably'.) In a much more limited fashion, Chinese focus particle *shi* can also, as a suffix, optionally or obligatorily combine with some conjunctions and sentence adverbs, as shown in the last two examples of (56) (cf. Chao 1968: 722, Lü 1985, Zhang 2004). Further morphological and morphosyntactic studies are required to better understand the nature of these morphemes.

syntactically in that the former cannot license ellipsis or VP-preposing, while the latter can.[33]

(57) Will John go?

    a. He will.

    b. *He **certainly/obviously/fortunately**.[34]

(58) [Will he go?]

    a. ta hui

      he will

    b. *ta **yiding/dagai/xianran**

      he certainly/probably/obviously

(59) a. Love Mary, John does.

    b. *Loves Mary, John **certainly**.[35]

2.3.1.4 Can be coordinated with other adverbial expressions

(53d) is fairly restricted for sentence adverbs. It seems only evaluative adverbs can be coordinated. This is presumably due to semantic factors.

---

[33] Some Chinese dialects, however, allow subjects, adverbs and auxiliary verbs alike to be stranded in 'right-dislocation' constructions, as illustrated in the following example (Lu 1980, Cheung 2009, a.o.):

(i) dao jia   le ba, ta dagai   (Beijing Mandarin)
    arrive home Asp Prt he probably
    'He has probably got home now.'
(ii) zhen gao a, zhe lou!
    real   tall Prt this building
    'This building is real tall!'
(ii) nei   maai-zo   matje aa3, doudai?   (Cantonese)
    You buy-Asp what Prt   DOUDAI
    'What the hell did you buy?'

These cases do not seem to involve typical VP-preposing. I leave them for future research.

[34] Examples in (57b) would be ok if the subject is also elided. This seems to suggest they can license TP-ellipsis in English. Similar cases can be sometimes found with the complementizer *if*:

(i)   A: Please don't cheat.
      B: As if.
(ii) A: If he didn't find it, how would we find it?
      B: But what if?
(iii) He won't be here anytime soon, if ever.

The same cannot be said for Chinese examples in (58b), however, unless a sentence-final particle is added. Details aside, it is clear sentence adverbs have distinct PF-related properties compared to auxiliary verbs.

[35] In certain Chinese dialects, the counterpart of this sentence is possible. See note 33.

(60) a. **Unfortunately** and **sadly**, our country is plagued with a crisis.

    b. ***Unfortunately** and **definitely**, our country is plagued with a crisis.

    c. ***Definitely** or **probably**, our country is plagued with a crisis.

2.3.1.5 Generally do not select and are not selected

Property (53e) is widely acknowledged for sentence adverbs. Sentence adverbs are never selected arguments of a predicate, nor do they themselves select complements.[36]

2.3.1.6 Inflection marking is mostly absent

Property (53f) is also generally attested for sentence adverbs.[37] In English, adverbs in general never bear tense, aspect, and agreement marking (61a,b). Some adverbs accept comparative and superlative marking (61c), but not sentence adverbs.

(61) a. *John nevers like Bill.

    b. *John **certainlies** like Mary.

    c. John complained louder than Mary did.

    d. *John **simpler** likes Mary.

In Chinese, a morphosyntactic process called A-not-A construction, where the verb/auxiliary or the first syllable of the verb/auxiliary is followed by the negation marker *bu* and then the reduplicated form, can be regarded as an inflection marking strategy and can distinguish verbs or auxiliary verbs from adverbs and prepositions (Huang 1982). Sentence adverbs never take part in this morphosyntactic process.

(62) a. zhangsan xihuan lisi

    Z.   like   L.

    'Zhangsan likes Lisi.'

---

[36] Evaluative adverbs in English seem to be an exception. They can usually take a *for* PP complement (e.g. *Fortunately for you, he knows very little. Happily for the kids, the teacher didn't assign homework today*). This, however, is not in general possible for sentence adverbs in English and other languages.

[37] One notable exception to this property of adverbs is plural 'agreement' marking in Korean. According to Kim (1994), plural marking can appear everywhere in a sentence (but not on the sentence-final auxiliary verb), including on adverbs. A recent analysis in terms of Agree is provided by An (2007). An (p.c.) also observes that both VP adverbs and most sentence adverbs can be optionally marked. See Yim (2003) for slightly different judgments with regard to sentence adverbs. I will not go into details here.

b. zhangsan xi-bu-xihuan lisi?

    Z.   li-Neg-like   L.

    'Does Zhangsan like Lisi?'

(63) a. zhangsan **yiding**   xihuan lisi

    Z.   certainly   like   L.

    'Zhangsan certainly likes Lisi.'

    b. * zhangsan **yi-bu-yiding**   xihuan lisi?

      Z.   cer-Neg-certain like   L.

2.3.1.7 Do not change the category or bar-level of the constituent they are joined to

(53g) is a well-known property of all adverbial adjuncts, and sentence adverbs are no exception. Their inability to change the category of the attached constituents is illustrated as follows.

(64) a. John will talk to Mary and ignore Peter.

    b. John will talk to Mary and **probably** ignore Peter.

(65) a. *John plays piano doesn't focus.

    b. *John **definitely** plays piano doesn't focus.

(64) shows that a VP can be coordinated with a 'bare' VP (64a) as well a VP together with a sentence adverb (64b). (65a), intended to mean 'when John plays piano he doesn't focus', shows that the VP *plays piano* cannot occur in an adjoined position, and (65b) shows it still cannot when a sentence adverb is adjoined to it. These examples easily show that sentence adverbs do not have any effect on the category of the constituents they are adjoined to.

2.3.1.8 Optionality

Property (53h) is usually attested for sentence adverbs. However, it has been observed in Hole (2004) and Sung (2007) that in some contexts certain mood- or focus-related adverbs are obligatory. Consider the following examples from these authors:

(66) a. zhiyao   xingqitian tianqi   hao, wo *(**jiu**)   qu pa   shan

      if   Sunday   weather good I   JIU   go   climb mountain

      'I go mountain-climbing on Sundays if the weather is fine.'

b. jie   yi   huan   qi,   wo   *(**ye**) bu   jie
   lend one return   seven I   YE Neg lend
   'Even if I can get seven times as much as I've lent back, I will not lend money.'

c. zhe xigua   *(**ke**) tian   le
   this watermelon   KE sweet   Prt
   'This watermelon is REALLY sweet.'

d. you   baimian   wo   *(**hai**)   xiang   chi   ne
   YOU plain.noodle   I   HAI   want   eat   Prt
   'If there is plain noodle, I will still want to eat it!'

To anticipate later discussions, these mood and focus-related adverbs are sentence adverbs under our definition. Their obligatory presence is not expected and seems to be a problem for analyzing them as adverbs. The matter will be discussed in detail and resolved in chapter 4.

2.3.1.9 Recursion

With regard to the recursion property (53i), sentence adverbs seem to be restricted. The number of sentence adverbs in a sentence generally is not more than two, and they generally cannot be adjacent to each other in English (Jackendoff 1972). This may be due to both semantic and phonological factors. Some examples are illustrated below (Cinque 1999):

(67) a. **Honestly** I am **unfortunately** unable to help you.
     b. **Fortunately**, he had **evidently** had his own opinion of the matter.
     c. **Clearly** John **probably** will quickly learn French perfectly.

More cross-linguistic evidence can also be found in Cinque (1999).

2.3.1.10 Can be left or right adjoined to the target in certain cases

(53j) is difficult to establish without a detailed discussion of the interaction between syntactic positions and possible corresponding semantic shifts of adverbs and a more solid theoretical footing. I will postpone this discussion until chapter 4.

2.3.1.11 More distant from the head than complements

(53k) can be easily attested from the evidence of various constituency tests in English and

Chinese. We can use auxiliaries to test whether sentence adverbs form a constituent with the lexical verb. The results tell us they do not, at least in this context. When there are two auxiliary verbs in a sentence, sentence adverbs cannot occur after the second auxiliary verb in English (68b). In Chinese, sentence adverbs always have to precede the auxiliary verb.

(68) a. John had **luckily** been leaving the office at the time.

    b. *John had been **luckily** leaving the office at the time.

(69) a. zhangsan **xianran**   hui   tan gangqin

       Z.   obviously able.to play piano

       'Zhangsan can obviously play piano.'

    b. *zhangsan hui   **xianran**   tan gangqin

        Z.   able.to obviously play piano

Further, we can use the VP movement, which also shows that sentence adverbs and lexical verbs do not form a constituent, at least in this context. The preposed VP cannot contain a sentence adverb.[38]

(70) a. Play piano, John **definitely** can.

    b. ***Definitely** play piano, John can.

(71) a. tan gangqin, zhangsan **yiding**   hui

      play piano   Z.   certainly able.to

    b. ***yiding** tan gangqin, zhangsan hui

2.3.1.12 Attachment at different categorical levels

As mentioned in 2.2.1.1, (53l) is a controversial property. In the current theories of generative grammar, bar-level is eliminated in favor of a derivation-oriented phrase structure. (53l) must thus be understood as: is it possible for a sentence adverb to merge with a verb or some functional head before the latter merge with its complements? Note that this property conflicts with (53k), according to which an adverb is attached *after* all the arguments of the verb are combined with the verb. Based on our observation in (68-71), it does seem (53k) holds and (53l) doesn't. However, as mentioned above, based on (14) and (15) and other relevant data, a number of linguists argue that (53l) does hold, for both VP adverbs and sentence adverbs,

---

[38] This does not mean sentence cannot adjoin to VP in any context, however. The facts are compatible with an analysis according to which VP-preposing somehow 'bleeds' the possibility of VP-adjunction of sentence adverbs. I will return to this matter in the next chapter.

especially in Romance languages. So far syntactic theories have offered little to resolve this conflict, and mainstream theories generally assume that (53k) holds and (53l) doesn't without convincing arguments. As for proponents of (53l), they also have difficulties predicting when adverbs attach to an XP and when they attach to an X^0 element. I will return to this matter in chapter four.

2.3.1.13 Free word order in certain cases

(53m) is generally argued not to hold for sentence adverbs (cf. Cinque 1999 and references cited there). However, Nilsen (2001) and Ernst (2002) note that sentence adverbs can either precede or follow certain classes of adverbs without apparent semantic differences.

(72) a. Ståle hadde **muligens alltid** spist neon   andres   hvetekaker (Norwegian)
       S.   had   possibly   always   eaten   somebody else   wheaties
       'Stanley had possibly always eaten somebody else's wheaties.'
    b. Ståle hadde **alltid muligens** spist neon   andres   hvetekaker
       S.   had   always possibly   eaten   somebody else   wheaties
       'Stanley always possibly ate somebody else's wheaties.'
(73) a. Management will **therefore hardly** be ready to offer a new contract.
    b. Management will **hardly therefore** be ready to offer a new contract.
(74) a. We are **still probably** north of Princeton.
    b. We are **probably still** north of Princeton.
(75) a. tamen **shenzhi shuobuding** hui   qu   xinjiapo   (Chinese)
       they   even   maybe   will   go   Singapore
       'They even maybe will go to Singapore.'
    b. tamen **shuobuding shenzhi** hui   qu xinjiapo
       they   maybe   even   will   go   Singapore
       'They maybe even will go to Singapore.'

Thus sentence adverbs do show some traits of the free word order property that is considered to be typical for adjuncts.

2.3.1.14 Apparent counter-cyclicity

(53n) is usually discussed in the context of examples like (32), repeated below.

(76) Which picture of Bill$_i$ that John$_j$ liked did he$_{*i/j}$ buy?

Lebeaux's (1988) original analysis is that the adjunct clause *that John like* is merged after the whole *wh*-phrase is moved. This violates cyclicity and the extension condition, according to which Merge to α is always at the edge of α (Chomsky 2004), and involves 'backtracking' the earlier stages of the derivation. When we consider the syntactic distribution of sentence adverbs, we also found they seem to be counter-cyclic. This can be illustrated in the following sentence:

(77) a. Mary **obviously** writes well.

   b. John likes **probably** most people in this class.

If we follow the standard assumption of syntax-semantics interface and that sentence adverbs take sentential scope, sentence adverbs should not enter the derivation of a sentence until the whole TP is constructed. What (77) shows, however, is that sentence adverbs seem to be merged to terms contained in TP, if we do not consider the ill-motivated possibilities of subject and verb movements in these cases. The merge of sentence adverbs thus seems to be apparently counter-cyclic.

2.3.1.15 Do not block agreement

(53o) is a property that can readily distinguish adverbs from verbs or auxiliary verbs. Auxiliary verbs are known to block agreements on the lexical verb in many languages[39], but adverbs do not have this property. This property is illustrated below:

(78) a. John can play guitar.

   b. John **certainly** plays guitar.

Sentence adverbs behave just like typical adjuncts in this respect.[40]

   We can thus conclude that sentence adverbs do have the core properties of typical adverbial adjuncts.

---

[39] According to Baker (2008: 212), however, a healthy minority of languages have agreement on the main verb as well as the auxiliary in at least some cases, and at least one language, Tzotzil, allows agreement only on the main verb in a sentence.

[40] The adjacency requirement on accusative case assignment in English (e.g. *John speaks fluently English. John speaks often to Mary. *For apparently Bob to be sick would worry Harriet.*) seems to be a counterexample to (53o), since case marking is treated as an Agree operation in recent theories (Chomsky 2000, 2001, a.o.). To account for this fact, we have to either propose a somewhat different Agree process for case marking, or argue that adjacency requirement is a result of something other than accusative case marking. I will leave this issue aside.

2.3.2 Sentence adverbs have properties of C^0 elements

In this section I will show that sentence adverbs have certain core properties of C^0 elements. This part of the definition is entirely new, and I will therefore mention a few potential problems before moving on. At first glance, (52b) seems paradoxical. How can an expression have properties of adverbial adjuncts and C^0 elements at the same time? If it has the *defining* properties of both adverbial adjuncts and C^0 elements, an immediate consequence is that a sentence adverb is both an AdvP and a C^0 at the same time. This is certainly not acceptable in any version of generative grammar. One the other hand, if we only require that it has defining properties of adverbial adjuncts but only some 'non-defining' properties of C^0 elements, then the question goes to why sentence adverbs have some non-defining properties of C^0 elements, and how they can be accounted for in theoretical terms. With these problems in mind, I will go on to provide evidence and also motivation for this definition, arguing that it has solid empirical support and thus paving the way for theoretical solutions in chapter 4.

I now list the properties of sentence adverbs that are also shared by C^0 elements.

(79) a. Ability to scope over the sentence subject.
    b. Restricted when under the scope of a clausemate C^0 element.
    c. Selection restrictions with V^0.
    d. Restricted in embedded clauses in other contexts.
    e. Clause-linking function.
    f. Denotation focus and quantification are usually not possible.
    g. Long-distance movement is not possible.

I will examine these properties one by one.

2.3.2.1 Ability to scope over the sentence subject

A C^0 element in principle takes scope over all materials in the IP that it c-commands.[41] Although neither a necessary nor a sufficient condition, this property can provide a hint as to whether an expression is a C^0 element or not.

In the literature, this test has been used to determine whether an expression is a C^0 element when different expressions have different scope properties. Collins (1991) uses this test to

---

[41] This statement has to be qualified by 'in principle' because sentential operators sometimes can take narrow scope with QNP subjects, as shown in (80a, 81b, 82a). I will argue in chapter 3 that this is because the complex semantics involving these operators allow them to take two kinds of scope at the same time.

support his proposal that *how come* is a C^0 element, whereas *why* is adjoined to TP or VP. Consider the following sentences:

(80) a. **Why** did everybody hate John?

     b. **How come** everybody hates John?

In the above examples, (80a) is ambiguous, having a 'family of questions' interpretation as well as a single-question interpretation, while (80b) can only have a single-question interpretation. According to Collins, this is because *why* binds a trace that is lower than the surface position of *everybody*, whereas *how come* is base-generated in the C^0 position.

    Tsai (2008a) offers a similar argument for the existence of analogous C^0 elements in Chinese. His examples are as follows:

(81) a. (nimen,)   meigeren **zenme** hui dai   yi-ben   shu?
         you guys   everyone   how   will   bring one-Cl   book
         'How come everyone will bring one book?' (*wh* wide scope)

     b. (nimen,)   meigeren **weishenme** hui dai   yi-ben   shu?
         You guys   everyone   why   will   bring one-Cl   book
         'Why will everyone bring one book?'   (ambiguous)

Similar to their English counterparts, (81a) has only a single-question interpretation while (81b) is ambiguous. According to Tsai, this shows that *zenme* to be an Int(errogative)^0 element (part of the split-CP system), whereas *weishenme* occupies a lower position.[42]

    When it comes to adverbs in declarative clauses, similar contrasts can be found, which therefore suggests sentence adverbs behave like C^0 elements. This was illustrated in (4), repeated below:

(82) a. 60 percent of the voters **probably** prefer Dole to Gore.   (ambiguous)

     b. 60 percent of the voters **intentionally** left their ballots blank. (QNO subj. > Adv)

     c. 60 percent of the voters **completely** reject Dole.   (QNO subj. > Adv)

Comparing these examples to Collins's and Tsai's interrogative data, we find a somewhat similar, somewhat different pattern: the similarity is that sentence adverbs have different scope

---

[42] Although I agree that *zenme* has properties of a C^0 element, my analysis will be different from Tsai's, since sentence adverbs such as *zenme* do not overtly occupy a C^0 position, and subjects here are not necessarily topicalized, as I will discuss in more detail in chapter 4.

possibilities than other adverbs, the difference here is that (82a) is ambiguous, whereas the previous examples argued to involve C^0 elements are not ambiguous. These facts are surprising for Collins and Tsai in that they do not parallel their analyses of *why* and *how come*. According to them, if an adverb is ambiguous with regard to its scope relationship with a QNP subject, the former must be adjoined to VP as its base position. Therefore, when the adverb undergoes A′-movement, it can either take scope at its landing site or reconstruct to its base position. In (82a), however, there is no such A′-movement, yet we still have scope ambiguity. It is not clear for them how this can be explained by simply saying that *probably* is a TP- or VP-adjunct.[43]

A plausible, yet unexplored analysis is that *probably* is a C^0 element that is phonologically realized either as a (right-adjoined) DP-adjunct or a VP-adjunct. (82a) is ambiguous because the C^0 can 'associate' either with the subject or the verb phrase. *How come* and *zenme* are also C^0 elements, but have different 'association' properties, so they do not induce ambiguities. A more detailed analysis along this line will be provided in chapter 4. The crucial point here is that if sentence adverbs such as *probably* are not C^0 elements, one will most likely predict they can only take narrow scope in (82a), contrary the fact.

2.3.2.2 Restricted when under the scope of a clausemate C^0 element

A typical C^0 element is restricted when under the scope of another C^0 element. This property seems to be both a necessary and sufficient condition for a C^0 element. That is, all and only C^0 elements are restricted when another C^0 element occurs in the same sentence.[44] The following examples illustrate this fact:

(83) a. *If does John like Mary, …
    b. *If never did John run so fast…
    c. *Had if only you told me earlier! (cf. If only you'd told me earlier!)
    d. *Should that it have come to this! (cf. That it should have come to this!)
    e. *If John leave tomorrow… (cf. I demand that John leave tomorrow.)

---

[43] A reconstruction-of-the-subject-NP approach has been proposed by Bayer (1996) to deal with the focusing adverb *even*, and by von Fintel and Iatridou (2003) to deal with sentences with modal auxiliaries. This approach has never been fully developed, however, and it is not clear how it can account for the lack of ambiguity with *only* and adverbs in Chinese in general. Adopting this approach therefore also doesn't help a TP- or VP-adjunct analysis of sentence adverbs.

[44] Although there also cases where (non)finiteness is sensitive to the type of complementizer of a sentence, it is inclusive whether the complementizer is actually a reflex of tense (cf. Pesetsky and Torrego 2001) or whether tense is a reflex of the complementizer (Chomsky 2008). I will abstract away from these considerations, since they do not bear on the main point of this study.

(83a-d) shows that a clause can not have two overt clause-typing C^0 elements in English. (83e) shows that subjunctive mood marking in English is constrained by the clause-typing C^0 element of that clause. Non-C^0 elements are typically not restricted.

(84) a. If John is singing a song…
      b. If John has sung a song…
      c. If John is admired by Mary…
      d. If John is arriving…
      e. If John ate my lunch yesterday…

(84) shows that the choices of aspects, voices, verb types, and various non-C^0 elements alike are not restricted when they are under the scope of a C^0 element. We can thus conclude property (79b) holds for C^0 elements.

It is well-known that sentence adverbs in general are restricted when they are under the scope of a C_0 element, as illustrated below:

(85) a. *Has John **surprisingly** arrived?
      b. *Never did John **probably** run so fast.
      c. *If it **probably** rains you may get wet.
      d. *If only I **probably** have a car!

The parallelisms between (83) and (84) suggest that sentence adverbs are C^0 elements. If they are not, it is unclear how the parallelisms can be accounted for.

There are also a number of cases where sentence adverbs can occur in the presence of other C^0 elements, as illustrated in the following examples:

(86) a. (I've made my position quite clear.) What could there **possibly** be to talk about?
      b. Have they not **mysteriously** been refusing to answer questions about the budget?
      c. If they hadn't **mysteriously** disappeared that day, no one would have noticed the missing funds.
      d. If you are **probably** going to leave soon, there's no point in getting a broadband connection at home.
      e. Has John **perhaps** been here before?
(87) a. zhangsan la-le   ma/ba/Ø   (Chinese)
      Z.   arrive-Asp Prt
      'Has Zhangsan arrived?/Probably Zhangsan has arrived./Zhangsan has arrived.'

b. zhangsan **nandao**   lai-le   ma/*ba/Ø

    Z.   can.it.be arrive-Asp Prt

    'Can it be that Zhangsan has arrived?'

c. zhangsan **jingran**   kandao lisi *ma/*ba/Ø

    Z.   surprisingly see   L.   Prt

    'I can't believe Zhangsan saw Lisi.'

d. zhangsan **yexu**   kandao lisi *ma/ba/Ø

    Z.   perhaps see   L.   Prt

    'Zhangsan perhaps saw Lisi.'

In (86a), the adverb *possibly* duplicates the function of the modal auxiliary *could*: expressing epistemic modality, while the sentence implies the possibility is quite low. (86b,c) are a negative question and a negative counterfactual conditional, respectively. The semantics of these sentences 'bring in the truth' of the propositions (Ernst 2009), and can therefore accommodate certain truth-value-sensitive sentence adverbs. In (86d), because the conditional 'serves to make manifest a privileged discourse context', and has the 'speaker-anchoring' property that typical conditionals do not have, it can accommodate sentence-adverbs (Haegeman 2006).[45] In (86e), *perhaps* carries along an implication that gives a suggestion as to a possible answer (Bellert 1977). The Chinese examples in (87) demonstrate similar facts. They show that sentence adverbs such as *nandao* 'can it be', *jingran* 'surprisingly', and *yexu* 'perhaps' can occur with C^0 elements, realized as sentence-final particles in Chinese, yet the co-occurrences are clearly restricted.[46] What we can conclude here for our present purpose is that complex semantic and syntactic factors at C^0 level can restrict or allow the 'special' occurrences of sentence adverbs. If sentence adverbs are not C^0 elements, it is not clear how to account for the special, shifted, functions of sentence adverbs in (86a,d), the 'unexpected' grammaticality of (86b,c), the restrictions in (86e),

---

[45] In fact, Haegeman's analyses of (85c) and (86d) refer to the presence of C^0 elements. According to her, the protasis in (85c) is a 'truncated' CP, while the one in (86d) is not. These analyses are unequivocally based on the assumption that sentence adverbs are C^0 elements. In Haegeman (2010), she discards the truncation analysis and adopts an intervention analysis. According to the latter, the presence of sentence adverbs blocks movements of operators such as *if* and *when* in the adverbial clauses. In this analysis, sentence adverbs are still treated as C^0 elements. I will not choose between these two analyses here. Note Haegeman also observes that examples like (86d) cannot be accounted for by approaches simply treating sentence adverbs as PPIs.

[46] It is controversial whether sentence-final particles in Chinese are C^0 elements (cf. Paul 2009) or C^0-related elements that occur in a 'special' position (cf. Biberauer et al. 2009). However, since either analysis involves the syntax of C^0 to some extent, it doesn't pose a problem for the present analysis. I also leave it open whether the exact relationships between them and sentence adverbs are morphological (they form a discontinuous constituent) or syntactic (they are separate C^0-related elements). In any case, it should be clear from (87) that some kind of syntactic dependency relations exist between those particles and sentence adverbs, and that the syntax of C^0, rather than a lower functional head, is involved.

and the co-variations in (87).[47] Therefore, these data can be regarded as further evidence that sentence adverbs are C^0 elements.[48]

Although much work remains to be done to deal with the syntax and semantics of co-occurring sentence adverbs and C^0 elements as illustrated above, the fact the co-occurrences between them, but not between C^0 elements and lower functional heads, are restricted suggests sentence adverbs are C^0 elements instead of lower functional categories.

2.3.2.3 Selection restrictions with V^0

In addition to overt clausemate C^0 elements, if a sentence adverb occurs in a clausal complement of a lexical verb, the verb can also act to restrict occurrences of sentence adverbs. Consider the following examples:

(88) a. They think that **actually** he was informed.

     b. *They demand that **actually** he should be informed.

(89) a. John thinks that Mary **probably/obviously/unfortunately** did not attend the meeting.

     b. *John regrets that Mary **probably/obviously/unfortunately** did not attend the meeting.

(90) a. ta yiwei lisi   **yiding**   xihuan zhangsan

       he think L.   definitely like   Z.

       'He thinks that Lisi definitely likes Zhangsan.'

     b. *ta zhidao lisi   **yiding**   xihuan zhangsan

       he know L.   definitely like   Z.

     c. * ta cai   lisi   **yiding**   xihuan zhangsan

       he guess L.   definitely like   Z.

     c. * ta xiwang   lisi **yiding**   xihuan zhangsan

       he hope   L.   definitely like   Z.

In (88a), the verb *think* selects a declarative, and therefore assertive, clause. *Demand*, on the

---

[47] The auxiliary verb *should* also acquires special functions when it co-occurs with C^0 elements or occur in certain embedded clauses, suggesting it is also a C^0-related element.

(i) Why should he have left? (deontic or epistemic reading)
(ii) We invited John too, lest he should feel left out.
(iii)If you should see him, please let me know.
(iv)It's surprising he should have been so late.

[48] Note that although a semantics-oriented PPI licensing approach can account for (86b,c), it is not clear how it can account for all the data in (86). Furthermore, even if this approach is on the right track, it has to be supplemented by a syntax-based theory that deals with various other facts.

other hand, selects a 'jussive' clause. The sentence adverb *actually* can only occur in the first type of clause. The verb *regret* in (89), according to Haegeman (2006), takes a 'presupposition' complement. Sentence adverbs cannot occur in this kind of clauses. Similarly, in (90), the verbs *yiwei* 'think', *zhidao* 'know', *cai* 'guess', and *xiwang* 'hope' select different types of clauses. Complements selected by *yiwei* should be an assertive clause, whereas complements selected by other verbs in (90) are not. And *yiding* 'definitely' can only occur in assertive clauses.

These examples show that even though an overt clausemate C^0 is not available or does not have effect on selection restrictions by itself, we can still observe the effect of clause-types on the occurrence of sentence adverbs in a subordinate clause. An analysis treating sentence adverbs as C^0 elements seems to be more plausible than alternative approaches.

2.3.2.4 Restricted in embedded clauses in other contexts

In addition to the constraints coming from clausemate C^0 elements and lexical verbs, we can also observe that clause-types manifested in other ways also have an impact on the occurrence of sentence adverbs. There are potentially many possible factors that can be involved, since there are many clause types that have distinct syntactic properties across-linguistically. Here I will discuss three such cases.

One significant factor is whether a sentence is a matrix clause or embedded clause. It has been observed by Sung (2007) and Irwin (2009) that certain sentence adverbs seem to be restricted to matrix clauses only.

(91) a. ni **ke** huilai le!
        you KE back Prt
        'NOW you are back!'
    b. *ta yiwei ni **ke** huilai le
        he think you KE back Prt
    c. *ta zhidao ni **ke** huilai le
        he know   you KE back   Prt
    d. ***ke** huilai de ren   hen duo
        KE back DE person very much
(92) a. Everyone is **SO** wearing gray this season.
    b. Mary **SO** aced that physics exam.
    c. *Jamie claims that everyone is **SO** wearing gray this season.
    d. *Jamie believes that everyone is **SO** wearing gray this season.
    e. *Jamie knows that Mary **SO** aced that physics exam.

f. *Jamie was surprised that Mary **SO** aced that physics exam.

g. *Jamie hates that Mary **SO** aced that physics exam.

According to Sung (2007)[49], the mood adverb *ke*, which has a intensifying function and also expressing the speaker's annoyance here, can only occur in the matrix clause. It cannot occur in embedded clauses in general. In (91b,c) it occurs in clausal complements of verbs, in (91d) it occurs in relative clauses, and the sentences are all unacceptable. Similarly, Irwin (2009) notes that what the sentence adverb *so* (she terms it 'drama *so*') in general cannot occur in embedded clauses, as illustrated in (92). This seems to suggest that the (covert) C^0 elements compatible with these sentence adverbs can in general only be found in matrix clauses.

Another factor is the realization of tense in the embedded clause. It has been noted by Taglicht (2001) that tense morphology of a verb form in English can affect the occurrence of the sentence adverb *actually*, as illustrated below:

(93) a. I hope that **actually** he won the game.

b. *I hope that **actually** he wins the game.

(94) a. *(He may stay on, but) if **actually** he leaves, we'll have to replace him.

b. If **actually** he's leaving us at the end of the week, we'll have to replace him.

According to Taglicht, the clauses in simple present tense are 'nonassertative', whereas other tenses manifest 'implicit positive bias' toward the truth value of the clauses. This seems to suggest that some covert C^0 element that governs the choice of tense also governs the occurrences of sentence adverbs such as *actually*.

Another factor is the function of the clause relative to the matrix clause. Clauses classified by this criterion include reason clauses, purpose clauses, result clauses, time clauses, etc. When no overt C^0 elements mark these clauses, we can still detect the effect of C since certain types of sentence adverbs can occur in certain types of adjunct clauses, but not others, as illustrated below:

(95) a. John **fortunately** knowing the answer, I didn't fail the test.   (reason)

b. *John **certainly** knowing the answer, I didn't fail the test.

(96) a. She explained how she risked her life to quite **possibly** save it.   (purpose)

b. *She explained how she risked her life to **fortunately** save it.

---

[49] Only example (90a) is from Sung (2007). He doesn't provide examples that involve various types of embedded clauses.

These restrictions obviously do not come from overt C^0 elements, but plausibly from the presence of covert C^0 elements connected to the function of the embedded clauses.

In sum, various factors that seem to bear on the presence of non-overt C^0 elements in a clause affects determine the occurrences of sentence adverbs. They are thus further evidence that sentence adverbs are C^0 elements.

2.3.2.5 Clause-linking function

Certain C^0 elements have clause-linking functions, as illustrated below:

(97) a. **If** John wins the lottery, life will be easier for him.
    b. **Although** times are tough, he still keeps a positive outlook on life.
    c. **Unless** Mary performs well, she will not get the prize.
    d. **Because** he is sick, the project has to be postponed.

In the above examples, *if*, *although*, *unless*, and *because* are arguably C^0 elements (Emonds 1985). Their primary function is to indicate the function of the clause in relation to another clause that is syntactically and semantically connected to it. These expressions clearly have clause-linking functions, and are obviously C^0 elements rather than T^0 or lower functional heads.

When we turn our attention to adverbs, we find there are actually many adverbs of the same function in English and cross-linguistically.[50]

(98) a. He has never had the disease himself but he can **nevertheless** identify it.
    b. His son had been charged with importing illegal drugs; Ed had **therefore** decided to resign from the School Board.
    c. You **either** leave now or I'll call the police.
    d. **Not only** was the price very high, but the performance was also horrible.
    e. I could have gone there. **Only** I didn't.

(99) a. ta **suiran**   qu-guo meiguo, ta bu hui   jiang yingwe   (Chinese)
      he   although go-Asp US   he Neg able.to speak English
      'Although he has been to the US before, he can't speak English.'

---

[50] See also Li (2005), who observes that many sentence adverbs (his term is 'modal adverb') in Chinese have this function at discourse level. From this perspective, it seems clause-linking function is a universal property of sentence adverbs. Since the syntax of discourse-linking elements in general is not well-understood, I will not discuss these cases here, but they certainly merit future research.

b. ni **ruoguo** kan naben shu, wu jiu kan zheben shu

   you if   read that book I   JIU read this   book

   'If you read that book, I read this book.'

c. ni   qu,   (na) wo **jiu** qu

   you go   then I JIU go

   'If you go I will go there.'

d. ni   qu,   (na) wo **cai** qu

   you go   then I   CAI go

   'Only if you go will I go.'

e. ta hen gao, (er) ta baba **que**   hen ai

   he very tall Conj he father instead very short

   'He's very tall, but his father is very short.'

(100) a. John-i   **pilok**   Mary-lul   salangha-ciman,…[51]   (Korean)

   John-NOM CA   M.-ACC   love-though

   'Although John loves Mary,…'   (Chung 2004)

b. John-i   **manil** Mary-lul salangha-myen

   John-NOM CA   M.-ACC   love-if

   'If John loves Mary,…'

(101) Peter hat **weder** das Theorem verstanden noch konnte Maria dem Beweis folgen.

   P.   has neither the theorem understand nor could M.   the proof follow

   'Neither has Peter understood the theorem, nor could Maria follow the proof.'[52]

                                      (German, from Lechner 2000)

In all of the examples above, the functions of the underlined expressions are all specifically related to clause-linking. They express logical or discourse-related relationships between two propositions, not just a single proposition. Their functions are, therefore, just like the C^0 elements in (97), despite their lower surface syntactic position.[53] On the other hand, although literature on these expressions is less prominent than on other adverbs, it is clear that they are adverbial adjuncts rather than pure functional heads based on the criteria established in (53). I will discuss some of these adverbs in more detail in section 2.4.

---

[51] CA stands for 'correlative adverb'. It does not by itself have semantic content, but pairs with focusing particles or complementizers. For details, see Chung (2004).

[52] Some English speakers in fact do not like *neither* to take sentential scope here, so the English translation is not acceptable to them. This fact is also noted by Wurmbrand (2008).

[53] See also Hole (2004) for a similar view on adverbs like *jiu* and *cai*. On the other hand, Vries (2005) and Zhang (2008) both argue that adverbs like *either* origin from functional categories ($Dist^0$, Co^0, or X^0) that conjoin the CPs, not from the CPs themselves. These analyses cannot easily explain why conjunctives and the adverbs can co-occur.

2.3.2.6 Denotation focus and quantification are usually not possible

A property of C^0 elements is that they cannot bear *denotation focus*[54] and quantification, due to their syncategorematic nature. The following examples illustrate this point.

(102) a. A: Who did you see?

      B: I saw JOHN.

  b. A: What did you do to Mary?

      B: I ENRAGED her.

  c. A: #What's the nature of the fact that he finishes his homework?

      B: #IF he finishes his homework.

  d. A: Can he play?

      B: IF he finishes his homework.

When an NP or the lexical verb is focused, a natural context is they correspond to missing pieces of information of the addressee, as shown in the question-answer pairs in (102a,b). However, the same cannot be said if a C^0 element is focused, as shown in (102c,d). The stress on the complementizer does not correspond to a missing piece of information; instead, the entire proposition that combines with the clause-typer/force indicator *if* is the denotation focus.

    The same pattern holds with other complementizers:

(103) a. What DID you do?

  b. A: [David smells like a zombie.]   (German)

     B: Ich denke, DASS er ein zombie ist.

       I   think that   he a   zombie is

       'I think that he is (indeed) a zombie.'

In (103a), the phonetic focus is placed on the auxiliary that occupies the complementizer position. This kind of focus can be used in a number of contexts. According to Creswell (2000), it can be used when (i) reasking the assigned topic question, (ii) speaker should know the answer but doesn't, (iii) repetition of salient question, (iv) question is still unanswered, and (v) requesting the value of a missing property. It is clearly not the auxiliary *did* that is highlighted semantically, since it does not have any semantic significance in the first place. In the German example in (103b), phonetic stress on the complementizer of the embedded clause also serves to highlight

---

[54] See Krifka (2007) for a definition. The kind of focus covers both identificational focus and information focus defined in Kiss (1998).

the proposition and strengthen the assertive force of the embedded proposition, as suggested by the context and English translation (cf. Gutzmann 2009).

The resistance of C^0 to focus and quantification is further illustrated in (104).

(104) a. *Only/*often DOES John smoke?

   b. I will cook, only/*often IF he cleans up.[55]

   c. *I won't cook, only/often UNLESS he cleans up.

(104a) shows that a question as well as a question operator cannot be focused or quantified. (104b,c) show that although a conditional clause can sometimes be restricted by *only*, it can never be quantified.

The semantic nature of focus on complementizers is not entirely clear, but these examples all seem to involving speakers' emphasizing the whole propositions combined with the force or mood associated with the complementizer. I will therefore tentatively call it *force-mood focus*.[56] Despite the lack of a complete semantic and syntactic analyses, it should be clear that C^0 elements themselves do not bear denotation focus.[57] It should also be clear that whether a CP can be focused depends on the logical function that C^0 plays with regard to other clauses it has logical relations with.

When it comes to adverbs in general sentence adverbs in particular, we find a similar pattern. Non-sentence adverbs in general can be stressed to indicate denotation focus, but sentence adverbs cannot.

(105) A: How did John sing?

   B: He sang BEAUTIFULLY.

---

[55] The fact that *if*-clause can be restricted by *only* and *even* seems to be an accidental result of the logical nature of *if*. In other languages, different complementizers have to be used.

[56] A better-known related phenomenon is focus associated with accenting the polarity elements such as the auxiliary (or main verb in certain cases) or negation, which are typically not treated as C^0 elements. In the recent literature, this kind of focus is often called 'verum focus' (a term created by Höhle 1992), and is generally analyzed as focusing on the truth-value alone. (e.g. A: *John didn't go*. B: He DID go.) However, considerations of various possible contexts involved in this kind of focus suggest things other than truth-vales can be focused. This should be clear in the following examples from Huddleston and Pullum (2002: 98):

(i)   Kim's the one who DID make a donation.
(ii) He didn't win, but he DID come in the first half dozen.
(iii) I AM pleased you can join us.

None of these examples just emphasize the truth values of the relevant sentences. Instead, it is the propositions themselves and how they are used that is focused. Therefore, it seems that these constructions should also be regarded as involving force-mood focus, not verum focus.

[57] See Gutzmann (2009) for an analysis in Kaplan's (1999) hybrid semantics framework.

(106) A: How likely does Mary play piano?

   B: #She **DEFINITELY** plays piano.

(107) It was YESTERDAY/***CERTAINLY** that she saw Mary.

(108) *Not only **POSSIBLY** but even **PROBABLY** they ran out of fuel.

(109) a. *He only/often **POSSIBLY** likes Mary.

   b. *Only/often **INCREDIBLY**, John can read two books in a day.

   c. *It was cold yesterday, she only/often **THEREFORE** got a cold.

It has also been noted in the recent literature that sentence adverbs can be stressed in certain cases, but the semantic effect is that the entire proposition combined with the sentence adverb is focused.

(110) a. Does John **REALLY** drink?

   b. Who can we **POSSIBLY** call at this hour of the night?

   c. I'm **SO** going to ace that physics exam.

   d. Hétvégére   ″**feltétlenül** elolvad a hó[58]   (Hungarian)

      By the weekend definitely melt   the snow

      'There's no doubt, the snow will have been melted by the weekend.'

   e. ta **JINGRAN**   xiang qu!   (Chinese)

      he surprisingly want go

      'I caN'T believe he wants to go!'

In (110a), we see the stressed epistemic adverb *really* has a special function: it can be used when one wants to ask a positive question but with an epistemic bias (Romeo & Han 2004). A similar bias obtains with epistemic adverb *possibly* in (110b). In (110c), the degree epistemic adverb *so* is stressed and expresses the strong commitment of the speaker (Irwin 2009). In (110d), the stressed epistemic adverb *feltétlenül* 'definitely' expresses strong certainty (Egedi 2009). Similarly, in (110e), the stressed evaluative adverb *jingran* express strong disbelief and surprise of the speaker. Although much work is necessary to nail down the exact syntax and semantics of stressed sentence adverbs, it should be clear these examples involve *force-mood focus*, instead of denotation focus, and they are never associated with run-of-the-mill focus operators or quantificational adverbs.[59]

---

[58] I leave intact Egedi's (2009) notation (″) to express primary stress.

[59] Note that the PPI-account of sentence adverbs cannot account for these facts, since quantificational adverbs and focus operators are not NPI-licensing elements, and should not have a negative impact on PPIs in general.

(111) a. *Does John only/often **REALLY** drink?

 b. *Who can we only/often **POSSIBLY** call at this hour of the night?

 c. *I'm only **SO** going to ace that physics exam.

 d. *ta zhi/yizhi   **JINGRAN**   xiang qu!
     he   only/always surprisingly   want go

We thus have one more parallelism between sentence adverbs and C^0 elements.[60,61]

2.3.2.7 Long-distance movement is not possible

A property of C^0 elements discussed in Collins (1991) is that they cannot undergo *wh*-movement, as shown in the following examples:

(112) a. Why did John say Mary left?

 b. How come John said Mary left?

As we saw, example (112a) is ambiguous: *why* can either be associated with the embedded clause or the matrix clause. (112b) is not ambiguous: *how come* can only be associated with the matrix clause. According to Collins, the contrast comes from the fact that *why* is a TP or VP adjunct, whereas *how come* occupies the head of CP. And since a head is subject to the Head Movement Constraint or C^0 is generally frozen for head movement, *how come* can never move from a C^0 position to another C^0 position.

I agree that the contrast in (112) is due to the fact (112b) involves C^0 movement, but I think it is not due to a general ban on C^0 movement per se. If we look at the derivations more closely, we can see C^0 movement in (112b) is preempted by selection restriction. It is well-established

---

[60] For analyses that involve polarity focus and verum focus, see also Laka (1990), Culicover (1991), Holmberg (2001), and van Craenenbroeck (2004) all deal with related phenomena. I assume these cases are all parts of force-mood focus and involve syntactic operations at CP level. For some discussions of focused mood particles, see Egedi (2009: 126) and Gutzmann (2009).

[61] Ernst (2002: 369ff) discusses a number of cases where sentence adverbs follow quantificational adverbs, which may seem like counterexamples to the assumption sentence adverbs cannot be focused or quantified. They are not. Although his examples still lack extensive descriptive and theoretical accounts, as Ernst himself admits, it is clear none of them involve quantifying or focusing sentence adverbs themselves. Some of his examples are as follows:
(i) The Lewinsky affair will always unfortunately stain Clinton's tenure in office.
(ii) They have often quite curiously found themselves alone even in a crowded city.
(iii)We are still probably north of Princeton.

These examples actually involve some ideosyncractic syntactic properties of English. I will return to them in the next chapter.

that selection is distinct from movement in that selection is local.[62] In (112a), since *why* is not the Q operator at C^0, no selection relation exists between *why* and the matrix verb *say*, so *why* can occur in the embedded clause as well as the matrix clause. On the other hand, since *how come* is a C^0 element, it necessarily establishes a selection relation with the matrix verb *said* in (112b). It is presumable that being at C means *how come* is an interrogative operator bearing a [Q] feature. Therefore, any verb that selects a clause that begins with *how come* must be a [Q] feature-selecting verb.

(113) John wonders how come Mary left.
       [*u*Q]   [*i*Q]

In (112a,b), the verb *said* cannot be selecting an interrogative clause[63], since *why/how come* occurs at the left edge of the matrix clause and can only be associated with one interrogative clause, in this case the matrix clause.

(114) *How come$_i$ John said t_i Mary left?
              [*u*Decl.] [*i*Q]

As the feature specification in (114) shows, if *how come*, bearing an [*i*Q] feature, is merged in the C^0 position of the embedded clause, selection requirement of the verb *said* is not satisfied. Therefore, it is impossible for *how come* to have moved from the lower C to the higher C in (112b).

    When it comes to sentence adverbs that appear in interrogative clauses, we find they parallel the distribution of *how come*. As noted by Tsai (2008a), the counterpart of *how come* in Chinese is *zenme* in Chinese. It differs from the adverb *weishime* is that the former cannot

---

[62] There are cases of apparent counterexamples to the local nature of selection restriction. As discussed in Adger and Quer (2001), there are cases of *unselected embedded questions* (UEQs), where higher functional heads of the matrix can have an impact on the type of embedded clauses that can occur. Typical examples are as follows:

(i) *Julie admitted/heard/said if the bartender was happy.
(ii) Did Julie admit/hear/say if the bartender was happy?
(iii)Julie didn't admit/hear/say if the bartender was happy.

A&Q argues that these examples involve QR of the DP that contains the *f*-clause and PSI(Polarity-sensitive Item)-licensing, so the locality requirement of selection is still obeyed. I will not further discuss the issue here, since examples given here do not involve UEQs.

[63] As is well known, *say* behaves like *know* in that it can select declarative clauses or interrogative clauses.

(i) John said Mary was drunk.
(ii) John said who was drunk.

undergo (covert) long distance movement, while the latter can, as illustrated below:

(115) a. akiu   renwei xiaodi   **weishenme**   hui   cizhi?
      A.   think X.   why   will   resign
      'Why does Akiu think [Xiaodi will resign t]?'
    b. *akiu   renwei   xiaodi   **zenme** hui   chuli   zhe-jian   shi?
      A.   think   X.   how   will   handle this-Cl   matter
      '*How come Akiu thinks [t[Xiaodi will handle this matter]]?'

Based on Collins's (1991) analysis of the similar contrast between *how come* and *why* in English, Tsai argues that *zenme* differs from *weishenme* in that the former is a C^0 element, while the latter is a TP-adjunct. In our account, this would mean that *zenme* has a [Q] feature that is incompatible with the matrix verb *renwei* 'think' in (115b), so merge of *zenme* in the lower C^0 is impossible.[64] This is just like our account of cases in 2.3.2.3 above.

    The present account can be extended to interrogative non-*wh*-adverbs that nevertheless can only occur in interrogative clauses. One such adverb is *daodi*, which expresses the speaker's impatience to know the answer to a question. While it can occur in declarative sentences in certain dialects, it has very different meanings and should be treated as a different lexical item. Some sentences involving *daodi* are illustrated below:

(116) a. zhangsan **daodi**   kandao-le shei?
      Z.   DAODI see-Asp   who
      'Who the hell did Zhangsan see?'
    b. **daodi**   shei kandao-le lisi?
     DAODI who see-Asp   L.
     'Who the hell saw Lisi?'
    c. %zhangsan **daodi** kandao-le lisi
      'After all, Zhangsan saw Lisi/At last, Zhangsan saw Lisi.'

A basic question is how syntactic theories can account for the fact that interrogative adverbs like *daodi* must occur in interrogative clauses, since they are neither [Q] operators nor *wh*-words. If we treat *daodi* as a C^0 element that doesn't express interrogative force but must be compatible with it due to selection restriction, then we have a ready solution. In (116a,b), since *daodi* co-occurs with an interrogative C, the sentences are grammatical. In (116c), the sentence has a

---

[64] Of course, *zenme*, being an adverb, is not treated as a C^0 per se in our final analysis. The point here is that it behaves like a C^0 element. What this means theoretically will be clear in chapter 4.

declarative C, interrogative *daodi* cannot occur. This approach predicts that in embedded clauses the occurrence of interrogative *daodi* is conditioned by the choice of the matrix verb. This prediction is borne out:

(117) a. zhangsan xiangzhidao lisi **daodi** kandao-le shei
    Z.   wonder   L. DAODI see-Asp   who
    'Zhangsan wonders who the hell Lisi saw.'
   b. ?*zhangsan renwei lisi **daodi**   kandao-le shei   (with interrog. reading of *daodi*)
     Z.   think   L. DAODI see-Asp   who

In (116a), *daodi* can occur in the embedded clause since the matrix verb *xiangzhidao* selects an interrogative clause. The interrogative *daodi*, bearing a feature only compatible with an interrogative clause, is allowed to occur. In (116b), on the other hand, the matrix verb *renwei* selects a declarative clause, which does not satisfy the selection restriction of *daodi*. In this approach, there is no need to propose an ill-motivated long-distance movement account of interrogative adverbs (e.g. Huang and Ochi 2004), nor propose an unusual polarity item licensing account that fails to account for the strict locality (no clause-boundary crossing) requirement (viz. Law 2008).[65]

In sum, the ban on long-distance movement of various interrogative adverbs has a natural account if the latter are treated as C^0 elements, since selection is independently required and its well-established strictly local characteristic can account for all the relevant distribution facts once we know the feature make-up of the relevant C^0 elements. Under the present analysis, movement is preempted by selection restriction requirements (movement presupposes a lower merge position that will violate selection restricitons) and the lack of motivations (there is no feature to check or that relevant features are checked by selection only). It is doubtful all the facts can be properly accounted for if these adverbs are not treated as C^0 elements.

To sum up section 2.3.2, we have seen a number of facts that suggest sentence adverbs have properties shared by C^0 elements. They are able to scope over the subject. They are restricted under clausemate C^0 elements and when they are in embedded clauses. Certain sentence adverbs clearly have clause-linking as their primary function. They do not bear any kind of denotation focus. They do not undergo long-distance movement. All of these facts are unexpected if one simply assumes they are T^0 or lower functional head elements, nor can they be explained by

---

[65] Huang and Ochi (ibid) also have some descriptive oversights. According to them, (117b) should be grammatical, contrary to Law's and my intuition. If forced to give an interpretation to the sentence, the adverb *daodi* can never associate with the subject's attitude, but only with the speaker's attitude, which H&O also do not discuss since selection restriction is not a factor for them. See Law (ibid.) for some discussions of this issue.

semantics alone. Although formal analyses of all of the facts are still not transparent, partially due to the fact that CP-level syntax itself is still not well-understood, our new perspective should be an important step toward this goal.

2.3.3 Consequences

Although (52) is still a rudimentary, mostly pre-theoretical definition for sentence adverbs, it is able to achieve a considerable level of descriptive mileage while maintaining theoretical coherence. First, it is able to capture the time-tested intuition that sentence adverbs have a unique syntactic status, as has been discussed in 2.1. Second, by acknowledging that they are adverbial adjuncts, instead of functional heads or specifiers of functional heads, we can account for the fact that they generally behave as if they are 'not there apart from semantic interpretation', to use Chomsky's (2004) characterization of adjuncts, and have a freer distribution than non-adverbial adjuncts. Third, by explicitly making reference to C^0 elements, instead of referring to vague notions such as 'sentential scope' or 'speaker orientation', we are able to stay true to syntactic definitions and start to make sense of a wider range of their distributional facts that have been largely undiscussed in the literature.

2.4 On some non-typical cases

In addition to giving a more concrete definition to sentence adverbs, (52) also has some novel and welcoming consequences. Specifically, it covers a wider range of adverbs than the semantic notion 'speaker-oriented adverbs', which is a theoretical desirable result. It also sharpens the intuition that *mood* is a key ingredient of the syntax and semantics of sentence adverbs, without which certain important distinctions between different adverbs would be lost. In this section, I will illustrate these advantages by applying the diagnostics in 2.3 to certain adverbs and adverbial behaviors that have not seen much discussion in the literature. With (52) and the diagnostics discussed above, we can make sense of these adverbs that previous theories of adverb classification cannot.

2.4.1 The intensifying degree adverb *zhen*

Sung (2007) observes that there are a number of adverbs in Chinese that are semantically associated with strong subjectivity, and can only appear in the root clause. Here I will just focus

on the intensifying degree adverb *zhen*.[66]

The adverb *zhen* is typically understood as an adjective or adverb meaning 'real' and 'really', respectively. As an adverb, it usually is followed by a modification marker and can attach to all kinds of predicates and quantified subject NPs. When used as an intensifying degree adverb it can only attach to adjectives, some root modals, and psych verbs, and the modification marker *de* is always absent. I label the two *zhen*s as $zhen_1$ and $zhen_2$ respectively, for ease of exposition. Examples include the following:

(118) a. zhe shuiguo $zhen_1$-de hen tian
       this fruit   real-DE   very sweet
       'This fruit is truly very sweet.'

   b. ta $zhen_1$-de xihuan yuyanxue
      he real-DE like   linguistics
      'It's true he likes linguistics.'

   c. $zhen_1$-de meigeren dou   lai-le
      real-DE everyone DOU come-Asp
      'It's true that everyone came.'

(119) a. zhe shuiguo **$zhen_2$**   tian
       this fruit   ZHEN sweet
       'How sweet this fruit is!/This fruit is REALLY sweet!'

   b. ta **$zhen_2$** hui   chi
      he ZHEN able.to eat
      'He has such an appetite!'

   c. ta **$zhen_2$** xihuan   yuyenxue
      he ZHEN like   linguistics
      'Does he like linguistics or what!'

   d. *ta **$zhen_2$**   kandao-le lisi
       he ZHEN   see-Asp   L.

   e. ***$zhen_2$** meigeren dou lai-le
       ZHEN everyone DOU saw-Asp   L.

This adverb, unexpected for current theories of adverb classification, has many properties of typical sentence adverbs. For starters, a key feature of $zhen_2$, according to Sung, is that it can only occur in the root clause, as shown below:

---

[66] Other adverbs Sung discussed are *ke*, *ye*, *dao*, *dou*, *cai*, *jiu*, *you*, *zai*, and *hai*.

(120) a. ta **zhen₂** hui shuohua!

 he ZHEN able.to speak

 'Does he have a glib tongue or what!'

 b. yaoshi ta (***zhen₂**) hui shuohua, na laoban yiding xihuan ta

 if he ZHEN able.to speak then boss certainly like him

 'If he has a glib tongue, the boss will like him.'

As shown above, in a conditional clause, *zhen₂* cannot occur. Examples that involve other kinds of embedded clauses also show the same pattern:

(121) a. ta renshi yixie (***zhen₂**) xihuan yuyenxue de xuesheng

 3S know some ZHEN like linguistics DE student

 'He knows some students who like linguistics.'

 b. (***zhen₂**) xihuan yuyanxue shi jian hao shi

 ZHEN like linguistics be Cl good thing

 'Liking linguistics is a good thing.'

 c. yinwei ta (***zhen₂**) xihuan yuyenxue, suoyi mai-le henduo yuyenxue de shu

 because 3S ZHEN like linguistics so buy-Pft many linguistics DE book

 'Because he likes linguistics, he bought many books on linguistics.'

 d. ta yiwei lisi (***zhen₂**) xihuan yuyenxue

 3S think L. ZHEN like linguistics

 'He thinks Lisi likes linguistics.'

According to Sung, *zhen₂* expresses mood, and is specified with the feature [+main clause]. Its syntactic distributions all follow from this feature specification. This analysis, however, neither fits into current theories of adverb classification nor sufficiently account for two other distributional facts of this adverb, as we will see below.

Second, the distribution of this adverb is highly restricted under various C^0 elements, including clause-type, mood, and other sentence adverbs.

(122) a. ta (***zhen₂**) xihuan yuyanxue ma?

 3S ZHEN like linguistics Prt

 'Does he like linguistics?'

b. ta jingran/dagai/xianran (***zhen₂**) xihuan yuyanxue

    3S surprisingly/probably/obviously ZHEN like linguistics

    'Surprisingly/probably/obviously, he likes linguistics.'

c. ta **zhen₂** xihuan yuyenxhe a₁!

    3S ZHEN like linguistics Exc

    'Does he like linguistics or what!'

d. ta (***zhen₂**) xihuan yuyenxhe a₂/ba/ou![67]

    3S ZHEN like linguistics RF/SA/FW

    'He does like linguistics! (I told you so!)'

    'He likes linguistics, don't you agree?'

    'Let me tell you, he likes linguistics.'

(122a) shows that *zhen₂* cannot occur in a question. (122b) shows it cannot co-occur with various sentence adverbs. (122c,d) shows among non-interrogative-mood sentence-final particles, it can only co-occur with the one that expresses exclamative mood.

    Third, *zhen₂* cannot be quantified or be associated with a focusing operator in a sentence.

(123) a. ta bu/meiyou (***zhen₂**) xihuan yuyenxue

    he Neg ZHEN like linguistics

    'He didn't like linguistics.'

  b. meiyouren (***zhen₂**) xihuan yuyenxue

    nobody ZHEN like linguistics

    'Nobody likes linguistics.'

  c. meigeren dou (***zhen₂**) xihuan yuyenxue

    everyone DOU ZHEN like linguistics

    'Everyone likes linguistics.'

  d. ta de chengji tongchang dou hen/***zhen₂** hao

    he DE grade usually DOU very/ZHEN good

    'His grades are usually very good.'

The latter two properties cannot be account for by the [+main clause] feature as proposed by Sung. They also don't follow from the semantics of degree adverbs in general, since other degree

---

[67] The sentence-final particles in Chinese are difficult to translate. Here I follow Li and Thompson's (1981) translation. According to them, *a* can express 'impatient statement', which belongs to a more general category of 'reduced forcefulness', *ba* can express 'solicit agreement', and *ou* can express 'friendly warning'. The 'impatient statement' a_2 should be distinguished from the a_1 with a different stress pattern that expresses typical exclamative mood, not discussed by L&T.

adverbs such as *hen* can occur without problems.

(124) a. ta   renshi yixie   hen   xihuan yuyenxue de xuesheng
          he   know   some very like   linguistics DE student
          'He knows some students who like linguistics very much.'
       b. hen   xihuan yuyanxue shi jian hao   shi
          very like   linguistics be Cl   good thing
          'Liking linguistics is a good thing.'
       c. ta   hen   xihuan yuyenxue a_2/ba/ou!
          he very like   linguistics RF/SA/FW
          'He does like linguistics very much! (I told you so!)'
          'He likes linguistics very much, don't you agree?'
          'Let me tell you, he likes linguistics very much.'
       d. ta meiyou hen   xihuan yuyenxue
          he Neg   very like   linguistics
          'He didn't like linguistics very much.'
       e. meiyouren hen   xihuan yuyenxue
          nobody   very like   linguistics
          'Nobody likes linguistics very much.'
       f. meigeren dou   hen   xihuan yuyenxue
          everyone DOU very like   linguistics
          'Everyone likes linguistics very much.'

The full syntactic distribution of *zhen2* thus cannot be accounted for under Sung's analysis. Currently known semantic theories (projection rules, PPI theories) of adverbs do not seem to provide any insight on these distribution facts, either. However, our syntactic-based definition (52) easily captures all of these properties. In our account, *zhen2* is a sentence adverb, having the properties of (52). Since *zhen2* is a sentence adverb, it has properties of a C_0 element, it is natural that it has all the properties mentioned above. It is only the fact that it is additionally specified for degree modification that makes it stand out somewhat from the better-known sentence adverbs, but even the fact is not surprising once a more detailed analysis of sentence adverbs is provided in chapter 4.

2.4.2 The contrastive mood adverb *ke*

Another adverb that has not been classified as a sentence adverb but nevertheless manifests

86

all the core properties the latter is the Chinese contrastive mood adverb *ke*. The use of this adverb is illustrated in the following examples:[68]

(125) a. jing   che **ke** jiu zai qianmian
       police car KE right at front
       'A police car is right in front of you! (In case you haven't noticed.)'
  b. zhe yi ti   **ke** ba wo nan-zhu le
     this one problem.set   KE BA I   stump   Prt
     'I am stumped by this problem set! (And I thought I was smart.)'
  c. lisi **ke** mei shuo-guo zhe ju   hua
     L. KE Neg say-Exp this sentence word
     'Lisi didn't say this! (I'm telling you!)'
  d. zhe jian shi   **ke** bu   xunchang
     thisCL   matter KE Neg usual
     'This matter is unusual. (Other matters are quite common/Others think it's common)'
  e. ruguo ni   renwei zhangsan hen gao, ni **ke** jiu cuo   le
     if   you think   Z.   very tall you KE JIU wrong Prt
     'If you think Zhangsan is very tall, you are wrong. (I'm telling you!)'

As shown above, there is no direct translation of *ke* into English expressions; however, the interpretative effect it has on the sentence is quite consistent: the statement is in contrast to an assumption or another statement that is salient in the discourse or the common ground.[69]

According to current theories of adverb classification, it seems *ke* is, for all intents and purposes, a focusing adverb; since on the one hand, it is discourse-related, and on the other hand, it involves the semantics of information structure. However, simply treating it as a focusing adverb fails to capture many of its core properties, as we will see when we test it against (52).

First, the following properties indicate it is an adverb: (i) it can attach to various verbal, adjectival, preverbal, and pre-adjectival elements, so it's not a verbal affix (126);[70] (ii) it cannot be a stand-alone predicate, so it's not an auxiliary verb (127).

---

[68] *ke* as an adverb apparently has many semantic functions, and should perhaps be treated as homonyms. Its other functions include degree intensification, imperative mood intensification, expressing long-awaited wish-fulfilling, etc. (cf. Lü 1980, Luo and Shao 2006, Sheng 2006, a.o.). I will not discuss these uses here.

[69] In some cases, it can be translated as 'however' or 'but', but not in others. These translations correspond more directly to *raner* and *danshi*, respectively.

[70] However, *ke* cannot occur in the pre-subject NP position except when the subject NP is focus-marked. I will address this issue in the next chapter.

(126) a. lisi **ke** yizhi dou xiaozhong guojia
　　　 L. KE alwaysDOU loyal.to country
　　　 'Lisi has always been loyal to the country. (Others think he has questionable loyalty)'
　　 b. zhe yi ti **ke** ba wo wen-zhu le (=125b)
　　　 this one problem.set KEBA I stump Prt
　　　 'I am stumped by this problem set! (And I though I was smart.)'
　　 c. zhangsan **ke** conglai mei xue-guo fayu
　　　 ZhangsanKE all.alongNeg learn-Exp French
　　　 'Zhangsan has never learned French. (Others may think he has learned French)'
(127) wo mei xue-guo fayu, ta **ke** *(xue-guo)
　　 I Neg learn-Exp French 3S KE learn-Exp
　　 'Although I haven't learned any French, HE did.'

　　Second, the following properties show that it has properties of C^0 elements: (i) it's able to scope over the subject of the sentence (128), (ii) it's restricted when under the scope of a clausemate C^0 element (129), (iii) it is restricted in embedded clauses in general (130), (iv) denotation focus and quantification are not possible (131).

(128) a. meigeren **ke** dou xue-guo yingyu　　　　　　　　　　　　(*ke* > QNP)
　　　 everybody KEDOU learn-Exp English
　　　 'Everybody has learned some English. (Others think only some people have.)'
　　 b. **ke** meiyouren xue-guo fayu　　　　　　　　　(*ke* > QNP)
　　　 KEnobody learn-Exp French
　　　 'Nobody has learned any French. (Others think some people have.)'
(129) a. ruguo jing che (***ke***) jiu zai qianmian, na ni yao kai man yi dian
　　　 if police car KE right at front then you should drive slow a bit
　　　 'If a police car is right in front of us, you should slow down a bit.'
　　 b. zhe jian shi (***ke***) bu xunchang ma?
　　　 this matter KE Neg usual Prt
　　　 'Is this matter unusual?'
　　 c. zhe jian shi **ke** bu xunchang a_1/*a_2/ba/ou
　　　 this matter matter KENeg usual Exc/RF/SA/FW
　　　 'I'm telling you. This matter is unusual. (Other matters are quite common, etc)'
　　　 'This matter is unusual. (…) Don't you agree?'
　　　 'Let me tell you, this matter is unusual. (…)'

(130) a. lisi jiechu-le (*__ke__) ba wo nan-zhu de timu
   L. solve-Pfv KE BA I stump DE problem.set
   'Lisi solved the problem set that stumped me.'
  b. yinwei zhe-jian shi (*__ke__) bu xunchang, jingfang like zhankai-le diaocha
   because this-Cl matter KE Neg usual police right.away start-Pfv investigate
   'Because this matter is unusual, the police immediately start investigation.'
  c. zhangsan yiwei zhe-jian shi (*__ke__) bu xunchang
   Z.   think this-Cl matter KE Neg usual
   'Zhangsan thinks this matter is unusual.'
  d. zhangsan zhidao zhe-jian shi (*__ke__) bu xunchang
   Z.   know this-Cl matter KE Neg usual
   'Zhangsan knows this matter is not common.'
(131) a. zhexie timu (*dou) __ke__ ba wo nan-zhu le
   these problem DOU KE BA I stump Prt
   'These problem sets stumped me. (And I though I was smart.)'
  b. *meiyouren __ke__ bu hui shuo fayu
   nobody KE Neg will speak French
  c. laoshi (*changchang) __ke__ chu-le xie nan ti
   teacher often   KE set-Pft some tough problem.set
   'The teacher made up some tough problem sets. (Others thought they would be easy.)'

The above properties show that *ke* is not simply a focusing adverb or discourse-linking adverb, it has core properties of a sentence adverb. Other theories of adverb classification, such as those based on surface word order, truth-functional semantics, polarity-item licensing etc., are not able to account for all of these properties.

2.4.3 Adverbs that only occurs in certain non-declarative clause-types

Semantic and syntactic studies of sentence adverbs and speaker-oriented adverbs have often focused on adverbs used primarily in declarative sentences. This is because no theories of adverbs have predicted that sentence or speaker-oriented adverbs can only exist in certain non-declarative contexts. However, this kind of adverb abounds in Chinese, and has been noted in the descriptive literature (cf. Lü 1980). Such adverbs include *na*(*li*), *daodi*, *nandao*, *qianwan*, etc.

(132) a. ta **nail** shi guangdong ren? ta shi fujian ren. (rhetorical question)

3S since.when be Cantonese person 3S be Fujian person

'Since when is he a Cantonese? He is a Fujianese.'

b. *zhangsan **nail** shi guangdong ren. Lisi ye shi.

Z. since.when be Cantonese person L. also be

(133) a. lisi **daodi** zuo-le sheme? (question)

L. DAODI do-Pft what

'What on earth did Lisi do?'

b. %lisi **daodi** zuo-le henduo shi.[71]

L. after.all do-Pft many thing

'After all, Lisi did many things.'

(134) a. zhangsan **nandao** zou-le? (rhetorical yes-no question)

Z. can.it.be leave-Pft?

'Can it be that Zhangsan left?'

b. *zhangsan **nandao** zuo-le. lisi ye shi.

Z. can.it.be leave-Pft L also Aux

(135) a. ni **qianwan** yao xiaoxin! (imperative)

you QIANWAN must careful

'You really must be careful!'

b. *lisi **qianwan** hen xiaoxin

L. QIANWAN very careful

Let's again employ the diagnostics in 2.3.2 to test whether these adverbs are sentence adverbs according to our definition. Here, again, I will just choose one adverb, since the other adverbs behave similarly.

Let's examine the properties of *nandao*, which can be roughly translated as 'can it be that…' and expresses disbelief. It only occurs in (rhetorical) interrogative clauses which are marked by specific prosodic patterns. It has properties of adverbial adjuncts in that (i) it can occur in the sentence-initial position or attach to various pre-*v*P elements, so it's not a verbal prefix (136),[72] (ii) it cannot be stand-alone predicates, so it's not an auxiliary verb (137).

---

[71] As mentioned above, *daodi* can occur in declarative clauses for some, but not all speakers. I treat the declarative usage of *daodi* as a dialectal variation and a separate lexical item.

[72] In general, sentence adverbs, as well as focusing adverbs, cannot occur post-verbally except in the V-*de*-AdvP construction in Chinese. I will discuss this in more detail in Chapter 4.

(136) a. lisi zuotian zai gongyuan **nandao** pao-le yi tian?
L. yesterday at park can.it.be run-Asp one day
'Can it be that Lisi ran for a day at the park yesterday?'

b. lisizuotian **nandao** zai gongyuan yundong?
L. yesterday can.it.be at park exercise
'Can it be that Lisi exercised at the park yesterday?'

c. lisi **nandao** zuotian zai gongyuan yundong?
L. can.it.be yesterday at park exercise
'Can it be that Lisi exercised at the park yesterday?'

d. lisi **nandao** ba chuangci dapuo le?
L. can.it.be BA window break Pft
'Can it be Lisi broke the window?'

e. lisi **nandao** cong taibei chufa?
Lisi can.it.be from Taipei set.off
'Can it be Lisi set off from Taipei?'

f. **nandao** lisi zuotian zai gongyuan yundong?
can.it.be L. yesterday at park exercise
'Can it be that Lisi exercised at the park yesterday?'

(137) A: lisi hui qu gongyuan
L. will go park
'Lisi will go to the park.'

B: a. ta hui?
3S will
'He will?'

b. ta **nandao** *(hui)?

On the other hand, *nandao* also exhibit properties of C^0 elements in that (i) it is able to scope over the subject of the sentence (138); (ii) it is restricted when under the scope of a clausemate C^0 element; in (139), we see that *nandao* is only possible in yes-no questions that can take the particle *ma*; (iii) is restricted in embedded clauses in general (140); (iv) it can neither be the denotation focus nor be quantified (141).

(138) a. **nandao** meigeren dou qu-le?   (*nandao* > QNP)
can.it.be everyone DOU go-Pft
'Can it be that everyone went there?'

b. **nandao** youren qu-le? (*nandao* > QNP)

   can.it.be someone go-Pft

   'Can it be that someone went there?'

(139) a. zhangsan **nandao** xihuan lisi ma?

   Z. can.it.be like L. Prt

   'Can it be that Zhangsan likes Lisi?'

  b. zhangsan (***nandao**) xihuan shei ne?

   Z. can.it.be like who Prt

   'Who did Zhangsan like?'

  c. zhangsan (***nandao**) xi-bu-xihuan lisi ne?[73]

   Z. can.it.be li-Neg-like L. Prt

   'Does Zhangsan like Lisi?'

(140) a. zhangsan (***nandao**) xihuan lisi caiguai!

   Z. can.it.be like L. like.hell

  'Like hell Zhangsan likes Lisi! (lit. It'd be really strange if Zhangsan likes Lisi.)'

  b. yinwei tianqi (***nandao**) bu hao, suoyi jichang bixu guanbi

   because weather can.it.be Neg good so airport must close

   'Because the weather is bad, the airport must close.'

  c. ruguo zhangsan (***nandao**) qu, wo jiu go

   if Z. can.it.be go I JIU go

   'If Zhangsan goes, I go.'

(141) a. meigeren dou (??**nandao**) qu-le?[74] (*nandao* > QNP)

   everyone DOU can.it.be go-Pft

   'Did everyone go?'

  b. lisi changchang (***nandao**) qu meiguo ma?

   L. often can.it.be go US Prt

   'Does Lisi go to the US often?'

  c. lisi meiyou (***nandao**) qu ma?

   L. Neg can.it.be go Prt

   'Didn't Lisi go?'

---

[73] A-not-A questions per se are not incompatible with sentence adverbs. The adverb *daodi*, for example, can occur in an A-no-A question.

(i) ta daodi xi-bu-xihuan lisi ne?

  'Does he like Lisi or not?'

[74] For some speakers, *meigeren* as well as the distributor *dou* can precede *nandao*. This is akin to cases we discussed in n. 59, since *nandao* itself is not quantified, but something else is. I will not discuss this case further.

d. meiyouren (***nandao**)　　qu ma?

　nobody　　　can.it.be　　go Prt

　'Did nobody go?'

The above facts clearly show that there exist sentence adverbs that occur in specific non-declarative clause-types. The facts straightforwardly follow from (52), but not from semantic theories that simply treat speaker-oriented adverbs as predicates taking propositional objects as their arguments, nor from theories that treat speaker-oriented adverbs as PPI elements. The facts also cannot be accounted for straightforwardly in theories that treat sentence adverbs as VP or TP-level specifiers or adjuncts, which makes no reference to clause-types.

2.4.4 Connective adverbs

There are very few syntactic analyses of connective adverbs as a whole in the literature. In English, these adverbs include *moreover*, *alternatively*, *right*, *nevertheless*, *on the one hand*, *on the other hand*, *therefore*, *also*, *either*, *neither* and various that can be found in English grammar books (e.g. Huddleston and Pullum 2002). The lack of syntactic analysis is presumably due to the fact that they have conflicting syntactic properties. On the one hand, they behave like two-place predicates in that they generally require the existence of two clausal 'arguments'. On the other hand, they have clear properties of adverbial adjuncts. According to the classic analyses of adverbial adjuncts, they shouldn't have properties of augument-selecting predicates, these facts thus provide a challenge to the syntactic theories. Under the present definition of sentence adverbs, however, we can determine that connective adverbs, when they are involved in connect two clauses, are a type of sentence adverbs, and they don't need to be the predicates themselves, but the reflex of the selectional properties of the predicate, as has already been well-established in syntactic theories.

It is easy to see that connective adverbs fall into sentence adverbs as defined by (52). For expository reasons, I will focus on Chinese connective adverb *jiu*, which appears in certain conditional sentences and a number of other contexts.

Jiu as a connective adverb is illustrated as follows:[75]

---

[75] *Jiu* as an adverb has many other semantic functions, including a variety of temporal-related functions, emphatic functions, exclusive focus, scalar-focus, etc. See Lü (1980) and Hole (2004) for some discussions. I will only deal with its conditional use here.

(142) a. ruguo zhangsan qu, lisi **jiu** hui qu

       if   Z.   go L. JIU will go

      'If Zhangsan goes, then Lisi will go.'

    b. zhangsan yi   gaoxing **jiu** hui liao ge bu ting

      Z.   one happy   JIU will talk GE Neg stop

      'Whenever Zhangsan is happy, he talks nonstop.'

    c. meishi   **jiu**   duo zuo yihuir

      nothing JIU more sit a.while

      'If there is nothing you need to attend to, stay a while longer.'

In all of the above sentences, there is a conditional relationship between two clauses, where *jiu* clearly has a clause-connection function.

Is *jiu* a sentence adverb? It is not that it is from the current theories of adverbs. However, it is clear if we test its status against the definition outlined in (52).

Jiu has the core properties of adverbial adjuncts in that (i) it can attach to the verb as well as various preverbal elements, so it's not a verbal prefix (143); (ii) it cannot be a stand-alone predicate (144); (iii) it cannot under any inflectional marking, such as the one involved in A-not-A question (145); and (iv) it is optional (146)[76].

(143) a. ruoguo ta   mang, ta **jiu** zai gongsi   chi wancan

      if   3S busy 3S JIU at company eat dinner

      'If he is busy, he eats dinner at the company.'

    b. ruguo dou mei wenti,   women **jiu** ba dian   chufa

      if   DOU Neg problem we   JIU 8   o'clock set.off

      'If there is no problem, we set off at 8 o'clock.'

    c. ruguo dou mei wenti,   women **jiu** cong taibei   chufa

      if   DOU Neg problem we   JIU from Taipei   set.off

      'If there is no problem, we set off from Taipei.'

    d. ruguo you ren   weiguei   tingche, wo **jiu** ba che   tuo-cou

      if   YOU person law-break parking   I JIU BA vehicle tow-away

      'If someone parks illegally, I tow the vehicle away.'

(144) ruguo zhangsan qu, lisi **jiu** *(hui)

      if   Z.   go L. JIU   will

---

[76] Optionality is usually not an option when it comes to connective adverbs, but what is crucial here is that at VP, *v*P, and even TP level, nothing requires the obligatory presence of this adverb.

(145) *ruguo zhangsan qu, lisi **jiu-bu-jiu**   hui qu?

    if   Z.   go L. KIU-Neg-JIU will go

    (cf. ruguo zhangsan qu, lisi hui-bu-hui qu? 'If Zhangsan goes, then will Lisi go?')

(146) ruguo dou mei wenti,   women ba dian   chufa

    if   DOUNeg problem we   8 o'clock set.off

    'If there is no problem,

It also has the core properties of C^0 elements in that (i) it is able to scope over the subject of the sentence (147); (ii) it is restricted under non-conditional clausemate C^0 elements; in (148), we see that when *jiu* occurs in non-conditional contexts, it is either not legitimate or has a distinct focus-related meaning; (iii) it can neither be the denotation focus nor be under the scope of quantification (149).

(147) a. ruguo ni   bangmang, meigeren **jiu** dou hui zhichi ta   (*jiu* >QNP)

      if   you help   everyone JIU DOU will support him

      'If you help, then everyone will support him.'

    b. ruguo ni   fandui, na   **jiu** meiyouren hui zhichi   ta   (*jiu* >QNP)

      if   you oppose then JIU nobody   will support him

      'If you oppose, then nobody will support him.'

(148) a. ta **jiu** hui qu

      3S JIU will go

      'HE will go.'

    b. ta (***jiu***) hui qu ma?

      3S JIU will go Prt

    c. (***jiu***) likai!

       JIU leave

    d. ni   **jiu** likai ba!

      You JIU leave Prt

      'You should just leave!'

(149) a. ruguo ni   bangmang, meigeren (*dou) **jiu** hui zhichi ta

      if   you help   everybody   DOU JIU will support 3S

      'If you help, everyone will support him.'

    b. ruoguo you qian,   wo (*changchang) jiu **hui** qu meiguo

      if   have money I   often   JIU will go US

      'If I have money, I will go to the US.'

95

    c. *ruguo ni   bangmang, meiyouren **jiu** hui zhichi   ta

      if   you help   nobody   JIU will support him

  d. ruguo lisi gaoxing, ta (*bu) **jiu** qu gongyuan

    if   L. happy   3S Neg JIU go park

    'If Lisi is happy, he will go to the park.'

Those facts indicate *jiu* is a sentence adverb. These facts can also be more or less duplicated with various other connective adverbs. (52) thus achieves what alternative theories cannot. For theories that classify adverbs according to lexical semantics and compositional semantics, it is not clear how connective adverbs should be classified, since they are not predicates that take propositional arguments. It is also clear they cannot simply be analyzed as positive polarity items, since their occurrence is contingent on the existence of a conditional or an otherwise clausal-relation context. Treating them as a class of adverbs that are distinct from typical speaker-oriented adverbs also clearly misses the mark, since they have all the properties shared by other speaker-oriented adverbs. On the syntax side, it is also clear that theories that treat sentence adverbs as IP-level specifiers fail to account for why they can only occur in specific clausal-relational contexts; it is also clear that treating certain clausal connective adverbs (i.e. *either*) as extra-clausal functional heads (e.g. Co^0) that doubles as focusing adverbs (Hendriks 2002, Zhang 2008) does not capture the fact that they share many features with typical sentence adverbs we have seen above.

In sum, the above case studies of certain relatively obscure adverbs show that definition (52) is superior to other current theories of adverb classification since they fail to predict the existence or the essential properties of these adverbs. Although we are not able to scrutinize various other potential adverbs, the evidence we have seems pretty robust.

2.5 Conclusion

In this chapter I examined the question of 'sentence adverbs' as a distinct and coherent class of linguistic items. I showed that although there are strong empirical motivations for identifying such a class, its theoretical status in modern syntactic theory is quite shaky. After reviewing relevant literature, I proposed a definition of 'sentence adverb' based on core empirical facts and minimal theoretical machinery. This definition gains us considerable empirical mileage and retains theoretical coherence, which previous theories have had much trouble with. Case studies of certain adverbs not previously classified as sentence adverbs further illustrate the value of this definition.

Armed with this definition, we will proceed to explore further empirical and theoretical

issues in the definition itself and sentence adverbs in general. Many questions are still unanswered. For example, what is the syntactic distribution of sentence adverbs? What makes them different from other classes of adverbs? How is it that a C^0-related element has the properties of adverbial adjuncts? How is it that an adverbial adjunct has the properties of C^0 elements? How can we refine syntactic theories of adverbial adjuncts? To begin to address these questions, it's reasonable to start from the syntactic distribution of sentence adverbs, which I explore in the next chapter.

3. Focus-sensitivity of sentence adverbs

Although the definition provided in chapter two gives us a general idea of the syntactic properties of sentence adverbs, it does not supply details. Specifically, it does not tell us why sentence adverbs have those properties mentioned in chapter one: (i) the syntax-semantics mismatch problem, (ii) adverbial adjuncts, (iii) unique syntactic distributions, (iv) focus-sensitivity, (v) heterogeneity, and (vi) cross-linguistic variations. The next logical step is thus a detailed investigation of these properties one-by-one. In doing so we will see that there is good reason to concentrate on the focus-sensitivity property of sentence adverbs, since, as it turns out, most of the other properties are direct or indirect consequences of this.

Very little has so far been said in the literature about the focus-sensitivity of sentence adverbs, and less still about the formal syntactic analyses of this property. This oversight will be redressed in this chapter. Evidence can be found in languages and specific constructions where adverbs typically do not generally freely occur in various positions. In these constructions and languages, the positions of sentence adverbs are clearly affected by which part of the sentence is focused. Although not all sentence adverbs clearly exhibit this syntactic pattern, the generality of the pattern seems indisputable.

The consequences of these findings are far-reaching. Nothing in current syntactic theory predicts that sentence adverbs are focus-sensitive. Nothing in current theories of focus predicts that focus-sensitive expressions include sentence adverbs. Furthermore, the syntax of focus-sensitive adverbs has always been relegated to a secondary role in theories of adverbial syntax, and given at best a cursory analysis. Our findings show that linguistic theory should aim to capture these facts under a unified analysis, and it should provide an account of focusing adverbs.

In what follows I will first discuss what it means to be a focus-sensitive element, and list the basic fundamental syntactic properties associated with them. I then provide in section 2 the semantic and syntactic evidence for treating typical sentence adverbs as focusing adverbs. More

specifically, I show that the focus-sensitivity observed with typical focus-sensitive adverbs is also seen with sentence adverbs. As a consequence, a theory that does not treat the latter as focus-sensitive adverbs makes wrong empirical predictions. Section 3 concludes the chapter and addresses the consequences and outstanding issues.

3.1 What is focus-sensitivity?
3.1.1 Preliminary definition and types of focus/focus-sensitivity

I adopt the following general[1] definition of *focus* in Krifka (2007), which accepts the central claim, although not necessarily the exact proposal, of Alternative Semantics (Rooth 1985, 1992):

(1)   A property F of an expression α is a focus property iff F signals that alternatives of the denotation of (parts of) α are relevant for the interpretation of α.[2,3]

Similarly, I adopt his definition for *association with focus*:

(2)   Semantic operators whose interpretational effects depend on focus are associated with focus.

To see how these definitions help us in syntactic analyses of relevant sentences, let's take a look at some typical cases of focus and association with focus, as shown in (3) and (4). Focus-sensitive expressions are in bold face, while foci themselves are underlined (which are normally, but not always, marked by phonological prominence, see notes 3 and 7). Most of the examples come from Kawamura (2007), Beaver and Clark (2008), and works cited there:

(3)   a. John likes <u>Mary</u>.
        b. John <u>does</u> like Mary.
        c. It is <u>Mary</u> that John likes.
(4)   a. John likes **only** <u>Mary</u>.
        b. John **always** grades exams <u>in the morning</u>.
        c. Mary **seems** to have fed <u>Fido</u> Nutrapup.

---

[1] A more precise definition that involves specific syntactic configurations will be provided in (6a).
[2] Krifka's original definition also covers 'expression focus' which involve metalinguistic uses of focus. I will not discuss this kind of focus here.
[3] Note this definition says nothing about the prosody of a given expression. This works to our advantage since prosody only indirectly reflects the 'focushood' of a given expression. See also note 7.

d. Dogs **must** be <u>carried</u>.

e. **Every** ship passed through the lock <u>at night</u>.

Examples in (3) illustrate focus marking without overt focusing operators. (3a) is a typical case of what is generally termed 'information focus' (É Kiss 1998), natural as an answer to a wh-question such as *who does John like?* (3b) is a case of what is generally termed 'verum focus' (Höhle 1992), stressing the truth value or force of the sentence. (3c) is a case of what is generally termed 'identificational focus' or 'contrastive focus' (É Kiss 1998), which involves exhaustively identifying an entity from a set of alternatives. It is clear that in all of these cases, alternatives are relevant for the interpretation of the sentences. In (3a), the speaker singles out the denotation of *Mary* from an alternative set of individuals.[4] In (3b), the speaker singles out the positive truth-value of the sentence from an alternative set of truth-values. In (3c), the speaker singles out the denotation of *Mary* from an alternative set of individuals, and the former exhaustively satisfies the semantic requirement of the sentence. Examples in (4) all have overt focusing operators, whose interpretational effects depend on some focused expressions.[5] (4a) can be paraphrased as 'John likes no one from the set of contextually salient alternatives to Mary.' Here it is clear that neither alternatives of the denotation of the VP *like Mary* nor of the sentence *John likes Mary* are relevant for the interpretation of *only*, since the sentences cannot be paraphrased as 'the only property that John has is that he likes Mary, not some other property' or 'the only thing I know is that John likes Mary, not some other stuff'. Since alternatives of the denotation of the VP and the sentence are not relevant for the interpretation of the sentence, they are not the focus, according to definition (1). Similarly, (4b) can be paraphrased as 'Whenever John grades exams, he does so in the morning, not in the afternoon or in the evening.'[6] (4c) carries the inference 'Mary fed some dog Nutrapup', with the speaker's only having indirect evidence that Fido, among the alternatives, satisfies the semantic requirement. (4d) can be paraphrased as 'If dogs are present, then those dogs must be carried, and not be put on the floor, etc.' (4e) can be paraphrased as 'Every ship which passed through the lock did so at night, not at some other

---

[4] One might argue the VP *like Mary* and the whole sentence *John likes Mary* can also be considered as foci, according to definition (1). However, treating VP or the whole sentence as the focus cannot reflect the proper interpretation of (3a). This will be made clear when we discuss cases in (4).

[5] Most of these sentences are in fact ambiguous (cf. Jackendoff 1972). (4b), for example, can also mean 'It is in the morning that John always grades exams.' In this case *in the morning* is not associated with *always*, but some other covert semantic operator. I'll discuss sentences with multiple focus-sensitive operators in §3.1.2.3.

[6] This paraphrase does not entail that focus associated with *always* inherently receives exhaustive interpretation. For example, a mother scolding a child playing with carrots in her plate may say:

(i) Your sister always EATS vegetables.

This sentence does not exclude the possibility her sister plays with vegetables too. It can be roughly paraphrased as 'Whenever there are vegetables around your sister, among the things she can do with vegetables, she eats them.'

time.'

Now, we can see that syntax is somehow involved with regard to focus in these sentences. First, in (3), the focused expressions have special semantic properties, and are also marked by special prosody. These semantic and phonological properties are not intrinsic lexical properties of the lexical items. How do these properties arise? A plausible source is the output of narrow syntax, which may be read by the phonology and semantics component.[7] This is not unlike agreement, which also involves lexical items' features getting valued from a word-external source. Second, in (4), the interpretations of the focus-sensitive expressions crucially depend on a separate set of expressions in the same sentence, the foci. This is somewhat akin to properties of binding between two nominal expressions. The semantic relationships between them are certainly not word-internal, and are established after some syntactic relations are established first.

Clearly more needs to be said about the syntactic reflexes of focus and focus-sensitivity, and it is preferable if we have more than just phonological and semantic evidence. Fortunately, we do have more than just phonological and semantic evidence to show that focus-sensitivity is present in syntax. The evidence comes from the existence of focus-related movements and the syntactic distributions of a wide range of focus-sensitive adverbs. The latter are my concentration in this chapter, since they are a relatively uncharted territory. Before we discuss them, however, we should take a general look at the types focus and focus-sensitivity that have been identified in the literature.

Focus can be classified into two types, according to whether an overt focus-sensitive expression exists or not. When no overt focus-sensitive operator is present, as we have seen in (3), the focus is usually understood as *free focus* (Jacob 1983). It is this kind of focus that has attracted the bulk of attention in the syntactic literature so far (see e.g. Grewendorf 2005, Horvath 2007 for an overview of relevant literature), particularly in cases where overt syntactic displacements of focused expressions are involved, the landing sites are either CP-periphery or *v*P-periphery. I will not go into details of this kind of focus in this thesis. The examples in (4), also known as *bound focus*, are the ones that are crucial for the purpose of this chapter. So far, however, there have been few thorough and systematic analyses of bound focus as a linguistic phenomenon. It seems the best we have right now is a list such as the one provided by Beaver and Clark (2008) and works cited there:

---

[7] This does not mean, however, that whichever expressions with prosodic stress are the semantic foci or are syntactically active for focus-related operations. See Wagner (2006), Horvath (2007), Shyu (2010), note 12, and §3.1.2.2.2 for further discussion.

(5) **exclusives:** *only*, *just*, *merely*, . . .

    non-scalar additives: *too*, *also*, . . .

    scalar additives: *even*

    particularizers: *in particular*, *for example*, . . .

    intensives: *really*, *totally*, . . .

    quantificational adverbs: *always*, *usually*, . . .

    determiners: *many*, *most*, . . .

    sentential connectives: *because*, *since*, . . .

    counterfactuals: *if it were* . . .

    emotives: *regret*, *be glad*, . . .

    superlatives: *-est*

    negation: *not*, *no*, . . .

    generics: *Mice eat CHEESE.*

    yes-no question: *Did he do FIFTY push-ups?*

    aspectual adverbs: *still*, *already*

    minimizing downtoners: *kind of*, *barely*, *hardly*, . . .

    maximizing downtoners: *at most*, *at best*, *at a maximum*, . . .

    reason clauses: *because-clause*, . . .

    . . .

There has been no systematic study of the focus-sensitivity properties of all of these expressions in generative grammar in any language. There have only been substantial works in generative grammar that address certain syntactic properties of *only* and *even* (and sometimes *also*, *too*, and *always*) and their focus-sensitivity properties (Anderson 1972, Jackendoff 1972, Tancredi 1990a, 1990b, Longobardi 1991, Bayer 1996, Kayne 1998, Horvath 2007, Wagner 2009). Jackendoff discusses *merely*, *truly*, *simply*, *hardly*, etc. in passing, and Tenny (2000) briefly discusses *almost*, *nearly*, and *not*. Beaver and Clark (ibid.) confine themselves to semantic and pragmatic analyses.

    In sum, in this section we provided basic definitions for focus and focus-sensitivity. These definitions are couched in pre-theoretical and mostly non-syntactic terms, and a speaker's paraphrase intuitions, but they can help linguists identify focus-sensitive expressions of various types, which appear to have word-external phonological properties and semantic properties. According to modern versions of generative grammar, theses properties indicate focus-sensitivity is encoded in narrow syntax. If not, we would have the undesirable consequence that the phonological and semantic properties of certain expressions are not derived from syntax. However, one may argue that this is perhaps indeed the case and that generative grammar is wrong. In the next section, I will provide direct syntactic evidence to show that focus and

focus-sensitivity are indeed encoded in syntax.

3.1.2 Syntactic properties of focus-sensitivity

The above definitions of focus-sensitivity are not couched in purely syntactic terms. However, they do help us to identify certain expressions that have certain specific syntactic properties. In this section I will show what syntactic properties are relevant for these expressions.

3.1.2.1 Focus, host, and scope as essential ingredients

There are four basic components relevant for the syntax of focus-sensitive expressions (FSEs), the FSE itself, focus, host, and scope. I provide the following definition for the latter three expressions:

(6) a. The *focus* of an FSE is the expression whose denotation's substitution by alternatives is relevant for the interpretation of the FSE.
     b. The *host* of an FSE is the syntactic constituent it merges with.
     c. The *scope* of an FSE is the syntactic domain within which it has the ability to affect the interpretation of other expressions.

These definitions are mainly descriptive and pre-theoretic, and will be refined below. It is, however, important to note how all of these components are necessary in the syntax of an FSE. In what follows, the examples all contain the adverb *only*, because it is the best-known case as a focus-sensitive adverb and manifests all the syntactic properties clearly. I will discuss the syntax of other focusing adverbs in the next subsection.

The existence of focus and its relevance to the syntax of an FSE, can be seen from the following examples, in addition to the prosodic prominence:

(7) A: John saw Mary and Peter.
     B: No. (*Only) John (only) saw (only) <u>Mary</u>.
(8) A: John and Mary saw Peter.
     B: No. (Only) <u>John</u> (*only) saw (*only) Peter.

Both (7) and (8) are natural conversations where the focusing adverb *only* can be used. In (7), speaker B is responding to speaker A's assertion. Here *Mary* is the focus because speaker B wants to indicate if the denotation of *Mary* is substituted by the denotation of *Mary and Peter*,

the assertion is false. The syntactic effect of this focus is that the FSE can only occur in the preverbal position and the pre-object NP position. In (8), on the other hand, the subject NP *John* is the focus, because speaker B wants to indicate if the denotation of *John* is substituted by the denotation of *John and Mary*, the sentence is false. The syntactic effect of the placement of focus in the subject NP is shown by a different set of possible positions for the FSE *only*. This time it can only occur in the pre-subject position. From these examples it is clear that *focus* plays a role in the syntax of the FSE.[8]

The existence of host needs little justification, since all focus-sensitive expressions clearly merge with some expression. It is also easy to see that the host needs to be distinguished from focus. In (7), and note 8, we see the host need not be the focus: (7) shows *only* can attach to VP even though the object DP is the focus, the examples in note 8 shows that certain FSEs occur in a fixed position regardless the position of the focus.

The existence of scope is attested in the following examples:

(9) a. **Only** <u>John</u> ate any kale.
    b. *John ate any kale.

(10) a. Mary **only** said that <u>John</u> stole a cookie.
      'Mary didn't say of anyone but John that he stole a cookie.'
    b. Mary said that **only** <u>John</u> stole a cookie.[9]
      'Mary said that nobody but John stole a cookie.'

(11) We are required to study **only** <u>syntax</u>. (*only* > *require*, *require* > *only*)

(9) shows that although the host and focus of the FSE is the DP *John*, its scope is the whole sentence. Otherwise, the polarity item *any* wouldn't be licensed. (10a) and (10b) are minimal pairs: the FSE has the same focus in both cases, but different scopes. In (10a) the scope of *only* is

---

[8] The effect of focus on the syntax of an FSE is not always seen, however. For example, there is no difference between the position of the FSEs in the following examples:

(i) a. John likes <u>Bill</u> **the most**.
   b. <u>John</u> likes Bill **the most**.
(ii)a. Mary **seems** to have fed <u>Fido</u> Nutrapup.
   b. Mary **seems** to have fed Fido <u>Nutrapup</u>.

I assume that focus is still syntactically active here and undergoes covert movement. See §4.2 for further discussion.
[9] Note, however, that if *only* is attached to the object DP in a subjunctive embedded clause, wide scope is possible (Longobardi 1991, Kayne 1998):

(i) She has requested that they read **only** *<u>Aspects</u>*.

I will return to this complication in §3.1.2.2.3.

the entire sentence, while in (10b) the scope of *only* is the embedded clause. Similarly, on one reading of (11), the scope of *only* is the entire sentence, while the focus and the host is the object the embedded clause. These examples should suffice to show that the scope of an FSE is distinct from its focus and its host.

3.1.2.2 Syntactic relations between a focusing adverb, the focus, the host, and the scope

If a linguistic expression has syntactic dependency relations with three other syntactic constituents, an obvious next question is to ask what the dependency relations are. The issues are complex, and there are considerable lexical variation among FSEs and across languages, but some core facts seem to remain constant and should be identified. In what follows, I will deal with the type of FSE that is our major concern: focusing adverbs (FAs).[10]

3.1.2.2.1 Syntactic relations between an FA and its host

The most notable feature of this dependency is what is lacking: the lack of either θ-related relations or categorical selection between an FA and its host. This does not mean, however, they can freely merge everywhere. Let's take a closer look.

Free attachment, except TP
This is a well-known general property of FAs (Bayer 1996, 1999), illustrated below:

(12) a. John likes **only** [$_{DP}$ Mary].
      b. John sings **only** [$_{PP}$ in his house.]
      c. John **only** [$_{vP}$ likes Mary].
      d. Someone **only** [$_{vP}$ played a prank on someone].
      e. John **only** [$_{T'}$ can sing this song].
      f. **Only** [$_{CP}$ that John didn't bring any present] was surprising.
      g. *** Even/*only** [$_{TP}$ John likes Mary].
      h. **Only/*even** [$_{CP}$ John likes Mary].
       i. **Either** [$_{TP}$ John likes rice or beans].
       j. [$_{TP}$ John saw Bill], **even/too**.
       k. John [$_{T}$ can] **only** play piano.

---

[10] FSEs can also be verbs (see example (5)) or inflectional affixes (e.g. Cantonese *-dak* (Tang 2002)), Japanese *-dake*, and Korean *-man* (they all mean 'only')), which have somewhat different syntactic and morphosyntactic properties due to factors other than focus-marking. I will return to these cases in chapter 4.

The FA *only* attaches to DP in (12a), PP in (12b), *v*P in (12c,d), T′ in (12e), T^0 in (12k)[11], and CP in (12f,h). (12g) shows that left-attachment to TP is not allowed for *only* when the focus is within the TP. (12i) show that attachment to TP is possible for some other FAs, but it is perhaps not its first-merge position. Note also that the focus may or may not be the host, although in general the focus can be the host.

Maximal projection host

An FA's host generally must be a maximal projection, as shown in the following Chinese, English, and German examples:

(13) a. ta **zhi**   [VP wei zhangsan [VP xie-le   yi-ben shu]]

b. ??ta   [VP wei zhangsan **zhi**   [VP xie-le   yi-ben shu]]

he for Z.   only write-Pfv one-Cl book

'He **only** wrote <u>one book</u> for Zhangsan.'

(14) a. Some students smoke [PP <u>in</u> the classroom] **even**.   (From Bayer 1999)

b. *Some students smoke [P <u>in</u>] **even** the classroom.

(15) a. Peter$_i$ <u>küsste$_j$</u> Maria **nur** [VP t$_i$ t$_j$].   (From Büring and Hartmann 2001)

P.   kissed M.   only

'Peter only <u>kissed</u> Maria.'

b. *Peter **nur** [V <u>küsste</u>] Maria.

P.   only kissed M.

(16) a. …weil   man den Wagen **nur** [VP in   die Garage <u>fahren</u>] darf   (ibid.)

because one the car   only   into the garage drive   may

'…because you may only drive the car into the garage.'

b. *…weil   man den Wagen in   die Garage **nur** [V <u>fahren</u>] darf.

because one the car   into the garage only   drive   may

The examples in (13) show that when an FA is attach to a projection of the verb, it must attach to the maximal VP, not just a part of the VP. The contrast between (14a,b) shows that while PP can be the host of *even*, P^0 can never be. Similarly, the contrast between (15a,b) shows that while VP can be the host of *nur* 'only', V^0 cannot. Similar facts are shown in (16). If V^0 is a possible host

---

[11] In modern minimalist theories, the notion of bar-levels is eliminated, so the description involving adjunction to T′ should be updated. Similarly, there is no longer reason to ban right-adjunction to an X^0. See §2.2.1.1 for arguments for this type of analyses of sentences like (12k).

for *nur*, we will predict (16b) to be well-formed, contrary to the fact.[12] Note here that there are obvious exceptions to this generalization. If Williams and di Sciullo (1987), Radford, (1988), Sportiche (1988), Iatridou (1990), Williams (1994, 2000) are correct about their proposal that in certain cases X^0 can be a host for adjunction (see §2.2.1, and (12k)), we will have to qualify our ban on non-maximal hosts. Here I will follow Williams's (1994) proposal that lexical and auxiliary verbs in French and auxiliary verbs in English allow adjunction, while lexical verbs cannot be adjunction sites in English. (The facts of German are still not clear, however.) I will return to this issue in §4.2.4 and §4.3.6.

The direction of FA attachment is generally fairly consistent in a given construction in a given language. Generally, an FA is left-adjoined to its host in English and German.[13] Systematic exceptions exist: there are those cases we just talked about, where the host is an auxiliary verb or a lexical verb in French and an auxiliary verb in English, an FA is right-adjoined with no style shift at all. It seems reasonable to assume directionality is largely determined in the *phonological component*. I will return to this issue in chapter 4.

---

[12] Instead of a purely syntactic ban, there may be a deeper, interpretational reason for exclusion of hosts that are non-maximal. It could be that in all of these examples, it is actually the XPs themselves that are treated as the foci for the interpretational purposes, and not just the heads. Motivations for this line of reasoning comes from examples where the hosts are maximal projections but the sentences are still ill-formed:

(i) a. [DP LITTLE boys] only are permitted to use these chairs.
    b. *[AP LITTLE] only boys are permitted to use these chairs.
(ii)a. No [DP participation of YOUNG GIRL in the game] can they permit.
    b. *The participation of no [DP YOUNG GIRL] in the game can they permit.   (Adopted from Horvath 2006)

Here I only mark the prosody stress, represented by capitalization, instead of focus. The contrasts here cannot be explained away by the ban on non-maximal hosts. The issue stems from *which* maximal projection is the host. To anticipate discussion in §3.1.2.2.2, (b) sentences might be ruled out because the FAs do not c-command their foci, as their foci are actually the underlined expressions in (iii) and (iv), instead of the prosodically marked ones.

(iii) a. [DP LITTLE boys] only are permitted to use these chairs.
    b. *[AP LITTLE] only boys are permitted to use these chairs.
(iv) a. No [DP participation of YOUNG GIRL in the game] can they permit.
    b. *The participation of no [DP YOUNG GIRL] in the game can they permit.

It is possible that at least some contrasts in (14)-(16) can be accounted for in a similar fashion. That is, the maximal XPs are actually the foci. I will leave this possibility aside, pending a more thorough understanding of interpretational issues.

[13] However, just like a typical adverbial adjunct (see §2.1), in many cases an FA can also right-adjoin to its host, possibly with a style shift, as shown in the following examples:

(i) a. [DP Passengers] only are permitted on the platform. (Brennan 2007)
    b. [DP Anna's father] even was arrested.

3.1.2.2.2 Syntactic relations between an FA and its focus

A fundamental syntactic property of an FA is that its syntactic position is 'sensitive' to the syntactic focus or foci in a given sentence, as we have briefly sketched in §3.1.2.1. A more thorough investigation shows that there are in fact four pieces of evidence that indicate the 'sensitivity', i.e. the syntactic dependency between an FA and its focus or foci.

C-command
In general, FAs c-command their foci at overt syntax, whether they are attached to clausal projections (VP, TP, etc.) or non-clausal projections.[14] This generalization has been stated as the Principle of Lexical Association (PLA) (Tancredi 1990a, 1990b, Bayer 1999), which is perhaps the best-known syntactic principle that governs the syntactic distribution of focusing adverbs. The PLA covers the facts in (17), with some notable exceptions illustrated in (18):

(17) a. (Only) John (*only) saw (*only) Peter.
    b. (*Only) John (only) drank (*only) some coffee.
    c. Someone (only) played (*only) [DP a prank] on someone.[15]
(18) a. (Even) John will (even) play (*even) cello.
    b. In Saint Petersburg, (*always) officers (always) escort (*always) ballerinas.

In (17a), the adverb the focus is the subject DP *John*. The FA *only* can only attach to the subject. When it occurs at a position that doesn't c-command its focus, the sentence is ill-formed. Similarly, when the focus is the VP, the FA cannot attach to a constituent inside the VP, as shown in (17b). When the focus is the entire clause, then we see the FA also cannot attach to the object DP, although it can attach to the VP. The general pattern is quite clear: an FA c-commands its focus. Some exceptions are shown in (18). FAs such as *even* and *always* do not need to c-command its focus when they are attached to the auxiliary verb or the VP. Despite these exceptions, we still see they are focus-sensitive syntactically in that *even* cannot attach to the object DP when the subject DP is in focus. In addition, we will see later in this section and in §3.1.2.3.1 that some general locality conditions shown that an FA is still focus-sensitive in the auxiliary position.

Beyond these relatively simple cases, a number of facts that are usually regarded as 'island

---

[14] Japanese seems to be an exception to this generalization. According to Aoyagi (1998: 143), a focusing particle attach to a DP subject/object can associate with a category that dominates it. I will return to this exception in §4.3.4.9.
[15] The example is adopted from Jaeger & Wagner (2003).

effects' (see Bayer (1996, 1999)) or 'pied-piping' (Drubig 1994, Krifka 2006) can also be accounted for by this generalization. Consider the following examples:

(19) [A: At yesterday's party, there were two strangers. One man talked to Mary. The other man
    talked to Bill. But I think John knows both of them. B: No….]
    a. John knows only [the man who talked to MARY].
    b. *John knows the man who talked to only [MARY].[16]

The focus that is associated with *only* in (19) should be *the man who talked to Mary*, not just *Mary*. This is so because the sentences can only be paraphrased as (20a), not as (20b):

(20) a. The only man that John knows is the man who talked to Mary.
    b. Ok. We both know that John knows some stranger, and that stranger talked to someone
       we know. But I think that someone-we-know can only be Mary.

In other words, to the speaker it is alternatives to the denotation of the expression *the man who talked to Mary* that are relevant for the interpretation of *only*, not the alternatives to the denotation of *Mary*. According to our definition (6a), then, the whole complex NP is the focus of the FA *only*. The contrast between (19a) and (19b) can now be accounted for by our c-command condition and PLA. In (19a), the FA c-commands its focus. In (19b) it does not.[17] Note that with the wide-scope reading of *only*, changing the context to make *Mary* the focus still cannot salvage (19b). As we see in the paraphrase in (20b), to make *Mary* the focus of wide-scope *only* the speaker cannot use a complex noun phrase, but have to use several sentences to express his or her thoughts correctly.

Based on the same reasoning, we can also account for the following contrasts:

(21) a. ⌜[[ANNA's father] even] was arrested⌝.
    b. *⌜[[ANNA even]'s father] was arrested⌝.
(22) a. [DP LITTLE boys] only are permitted to use these chairs.   (=(i) in note 12)
    b. *[AP LITTLE] only boys are permitted to use these chairs.
(23) a. No [DP participation of YOUNG GIRL in the game] can they permit.   (=(ii) in note 12)
    b. *The participation of no [DP YOUNG GIRL] in the game can they permit.
(24) a. The store is closed only [PP on SUNDAY].
    b. *The store is closed on only [DP SUNDAY].

---

[16] This sentence is well-formed with a narrow-scope reading of *only*, but this is not relevant to our discussion here.
[17] This also shows prosody is not a sufficient condition for determining the focus of an FA. See note 3, 7, and 12.

(21b) and (22b) have been regard as violations of the left-branching effect (Bayer 1999). (23b) has been treated as an illicit case of pied-piping (Horvath 2006). (24b) has been treated as the effect of a ban on P-stranding (Rooth 1985). None of these analyses have considered the possibility that it is the larger constituents that are the foci. In light of our discussion above, however, it is reasonable to assume that proper interpretation of the FAs requires the denotations of the larger constituents to be counted as semantic units, rather than the smaller ones.[18,19] Therefore it seems the larger constituents are the foci in these sentences, and the (b) sentences are ill-formed because the c-command condition/PLA is violated.

These facts all suggest that syntactic dependency exists between an FA and its focus as defined in (6).

Adjacency

It has been shown that generally, when an FA c-commands its focus, the two cannot be separated by a constituent that is not part of the focus. The generalization is not as well-known as the c-command condition, but it is important and describes two types of facts (see Büring & Hartmann 2001, Jaeger & Wagner 2003, Reis 2005). In SVO or SOV languages, some FAs cannot occur in the sentence-initial position if the VP or a constituent within it is in focus, because the non-focus subject NP intervenes. We have already seen this in (12g) (although (12i,j) do not conform to this generalization). Similar facts are attested in V2 languages such as German. Second, in languages with OV word order or with robust pre-verbal modifiers and with FAs that can attach to various verbal projections, it is observed that FAs generally do not occur before non-focus preverbal arguments or modifiers. In other words, whenever one encounters a [XP [YP [V]]] sequence, the following patterns hold:

---

[18] A detailed interpretational account, unfortunately, cannot be provided here, due to the lack of current understanding of information structure and the fact paraphrase tests are not always clear. However, the fact that the larger constituents can be attached by FAs is significant and is easily compatible with our account. I will provide further evidence of my interpretation-based account and address some general issues of island effects in §4.3.4.4.

[19] It is worth noting that there seem to be exceptions of the c-command condition. When the focus contains an indefinite NP, the P-stranding effect is nullified (Taglicht 1984: 70, Kayne 1998: 155). Furthermore, according to my informants, (ia,b) can be synonymous, and so can (iia,b). The facts would also be unexpected by classical approaches based on subject-island effects. I will return to these exceptions in §4.3.4.9.

(i) [John visited several people, including Jessica, Mary and Phillip.]
    a. Only [John's visiting MARY] is good news.
    b. John's visiting only [MARY] is good news.
(ii) a. Even [pictures of MARY] are beautiful.
    b. Pictures of even [MARY] are beautiful.

(25) a. [FA [<u>XP</u> [<u>YP</u> [<u>V</u>]]]]

 b. [XP [FA [<u>YP</u> [<u>V</u>]]]]

 c. *[FA [XP [<u>YP</u> [<u>V</u>]]]]

Examples can be found in Chinese and German (German data are from Jaeger & Wagner 2003):

(26) lisi zuotian zuo-le sheme?   (Chinese)

 'What did Lisi do yesterday?'

 a. ta **zhi** <u>zai jia xie xiaoshuo</u>

 b. *ta <u>zai jia</u> **zhi**   <u>xie   xiaoshuo</u>.

   he at home only write novel

   'He only <u>wrote his novel at home</u>.'

(27) lisi zai jia zuo sheme?

 'What does Lisi do at home?'

 a. ta zai jia   **zhi**   <u>xie   xiaoshuo</u>.

 b. *ta **zhi** zai jia   <u>xie   xiaoshuo</u>.

   he only at home write novels

   'He only <u>writes novels</u> at home.'

(28) Warum hat Peter Marias Fahrrad umgdreht? Ich glaube,…   (German)

 'Why did Peter put Mary's bike upside-down? I think…

 a. *…dass **nur** Peter Maria <u>einen Streich spielen wollte</u>.

 b. *…dass Peter **nur** Maria <u>einen Streich spielen wollte</u>.

 c. …dass Peter Maria **nur** <u>einen Streich spielen wollte</u>.

 d. *…dass Peter Maria <u>einen Streich</u> **nur** <u>spielen wollte</u>.

    that P.   M.   a prank   only play   wanted

   …that Peter only <u>wanted to play a prank</u> on Mary.'

(29) Warum hat Peter Marias Fahrrad umgedreht? Ich glaube,…

 'Why did Peter put Mary's bike upside-down? I think…

 a. *…dass **nur** Peter Maria das Fahrrad <u>reparieren wollte</u>.

 b. *…das Peter **nur** Maria das Fahrrad <u>reparieren wollte</u>.

 c. *…das Peter Maria **nur** das Fahrrad <u>reparieren wollte</u>.

 d. …dass Peter Maria das Fahrrad **nur** <u>reparieren wollte</u>.

    that P   M.   the bike   only repair   wanted

   …that Peter only <u>wanted to repair</u> the bike for Mary.'

In the above examples, the FA must occur adjacent to the focus.

Nonetheless, we cannot simply say an FA must occur adjacent to its focus. There are some lexical and systematic exceptions. Certain adverbs, such as *either*, *even*, and *too*, can (left- or right-)adjoin to a TP with being adjacent to their foci, as shown in (12i,j). Furthermore, in a VO language like English, whenever an FA is attached to *v*P or T′ and the focus is the object DP, the verb and the auxiliary intervenes between the FA and the focus (eg. (12c,e,k)). In a multiple-clausal sentence, an FA is even further away from its focus, as shown below:

(30) a. John only knows that Mary believes that Peter loves <u>Jennifer</u>.
    b. The professor has even asked students to learn <u>Mohawk</u>.

Finally, as we have seen in (13)-(16), when a non-maximal projection is the focus or contains the focus, the FA cannot be adjacent to the former due to the ban on non-maximal hosts. These exceptions are obviously due to different grammatical principles at work, which we will address in §4.3.4. The general pattern is still very clear, the existence of the adjacency condition suggests some dependency between an FA and its focus.

Clausemateness

When an FA doesn't c-command its focus, the focus must be in the same minimal clause as the FA before the former undergoes further independently-motivated A'-movements (see Hoeksema and Zwarts 1991 for a similar observation). The relevant facts are illustrated below:

(31) a. *[John **even** went home] [although he hadn't met his <u>advisor</u>].
    b. *<u>Mary</u> thought [John'd **even** play cello].
    c. *In Saint Petersburg, <u>linguists</u> assume [officers **always** escort ballerinas].

Similar facts can be found in Chinese. Some FAs, such as *ye*$_a$ 'also' and the concord/agreement-like element *dou/ye*$_e$ of *lian* 'even', do not need to c-command their foci (32)[20]. These particles are typically attached to a clausal projection (33).[21] The foci and FAs cannot occur in two separate clauses (34).

(32) a. <u>zhangsan</u> **ye**$_a$ likai-le
        Z.   also leave-Pfv
        '<u>Zhangsan</u> also left.'

---

[20] See Hole (2004) for discussion.
[21] *dou/ye* do not form a constituent with the preceding focus. As is well-known, pre-verbal modifiers as well as TP-internal topics can intervene between the focus and *dou/ye*.

b. **lian** zhangsan **dou/ye_e** likai-le

    even Z.   DOU/YE leave-Pfv

    'Even Zhangsan left.'

(33) a. *ye_a zhangsan likai-le

      also Z.   leave-Pfv

b. *ye_a zhangsan xihuan lisi

      also Z.   like   L.

(34) a. *zhangsan yiwei [lisi **ye_a**   likai-le]

      Z.   think L. also   leave-Pfv

    'Intended meaning: Zhangsan also thinks Lisi left.

b. ***lian** zhangsan yiwei [lisi **dou/ye_e**   likai-le]

     even Z.   think L. DOU/YE leave-Pfv

    'Intended meaning: Even Zhangsan thinks Lisi left.'

c. *[suiran   lisi mei jiandao laoban], [ta **ye_a** huijia-le]

   although L. Neg see   boss   he also return.home-Pfv

    'Intended meaning: Although Lisi also didn't see the boss, he still went home.'

d. *[suiran   lisi **lian** laoban mei jiandao], [ta **dou/ye_e**   huijia-le][22]

   although L. even boss Neg see   he DOU/YE return.home-Pfv

    'Intended meaning: Although Lisi didn't even see the boss, he still went home.'

Note, however, that if an A′-movement triggers the focus to move to the matrix clause, the sentence is well formed:

(35) lisi_i, zhangsan yiwei [t_i **ye_a**   likai-le]

    L. Z.   think   also leave-Pfv

    'Lisi, Zhangsan thinks also left.'

Those facts show again that there is a syntactic dependency relation between an FA and the

---

[22] The concord/agreement elements *dou/ye* can associate with an entire clause, however:

(i) wulun   tianqi   duo   cha, ta **dou/ye**   hui qu shiyanshi

   no.matter weather how   bad he DOU/YE will go lab

   'No matter how bad the weather is, he (still) goes to the lab.'

In this sentence, *dou/ye* is a clause-linking particle, and does not seem to have the same function as the *dou/ye* in simple sentences. But even if we treat them as FAs, they still do not violate the clause-mate condition just mentioned, since unlike (34c,d), the focus is now the entire antecedent clause, presumably adjoined to the consequence clause ([CP YP [CP …*dou/ye*…]]). Therefore, they are still in the same minimal clause. I will return to these clause-linking adverbs in §3.2.2.6.

focus, even when the c-command condition is violated.

Multiple foci

When there is more than one focus associated with an FA, all the foci show some syntactic dependency to the FA. The following examples from Chinese are a case in point:

(36) A: zhangsan jintian zhaodao-le gongzuo. lisi zuotian zhaodao-le gongzuo.

'Zhangsan got a job today. Lisi got a job yesterday.'

B1: (bu.) ??<u>lisi jintian</u>   **ye$_a$** zhaodao gongzuo.[23]

no   L. today   also find-Pfv job

B2: (bu.) <u>lisi</u>   **ye$_a$**   <u>zai jintian</u>   zhaodao gongzuo.[24]

no L.   also at today   find   job

'No. <u>Lisi</u> also got a job <u>today</u>.'

(36) shows that an FA can be associated with two foci that are relevant to complex contextual environments. In (36B2), the single underline and the double underline mark the two foci that are associated with *ye*, respectively. The single-underlined focus indicates its similarity to a salient alternative in the discourse, Zhangsan. The double-underlined focus indicates the alternative time point (yesterday) mentioned in the context is incorrect. Crucially, the ill-formed (36B1) shows that the FA cannot occur lower than the double-underlined focus. This can be accounted for naturally if *ye$_a$* here is associated with *zai jintian* and therefore has to c-command it, based on the c-command condition mentioned above. (In English this is difficult to see since the FA *also* and the temporal adverb are at different sides of the *v*P.) Note here that the first focus, *Lisi*, is not c-commanded by *ye$_a$*. We expect it to obey the 'clausemate condition'. This is indeed the case:

(37) A: zhangsan cai lisi you liang-bu che. wangwu cai lisi you qi wu che.

'Zhangsan guesses that Lisi has 2 cars. Wangwu guesses Lisi has 7 cars.'

B1: (bu.) <u>wangwu</u> **ye$_a$**   cai [$_{CP}$ lisi you <u>liang-bu che</u>].

no W.   also guess L. have two-Cl   car

'No. <u>Wangwu</u> also guessed that Lisi have <u>2 cars</u>.'

B2: (bu.) *<u>wangwu</u>   cai [$_{CP}$ lisi **ye$_a$** you <u>liang-bu che</u>].

(37B2) is bad because it violates both the c-command condition and the clausemate condition.

---

[23] The sentence is ok as long as *jintian* is not the focus.

[24] That the preposition *zai* 'at' has to be inserted here is presumably due to the morphosyntactic requirement that *ye* cannot attach to a nominal constituent.

The first focus *Wangwu* is not c-commanded by *ye$_a$* and they are not clausemates.

Although the details still need to be worked out, it seems we have direct syntactic evidence that an FA may have multiple foci. Otherwise, it is not clear how the contrasts in (36) and (37) can be accounted for.[25]

This generalization will become important again when we address the syntax of clause-linking adverbs in §3.2.2.8.

3.1.2.2.3 Syntactic relations between an FA and its scope-taking position

Examples (9)-(11) have shown us the concept of scope is required for proper interpretation of an FA. In what follows I will provide four pieces of syntactic evidence further indicating syntactic dependencies between an FA and its scope-taking position.

Locality and ECP

It has been observed that when an FA attaches to a non-clausal constituent such as DP or PP, the resultant constituent [FA DP/PP] cannot freely occur in a sentence. Instead, it is subject to some locality constraints and displays ECP effects with regard to the scope position of the FA. The facts present perhaps the best-known argument for the covert movement of the [FA DP/PP] constituent (Longobardi 1991, Bayer 1996, Kayne 1998), and can be categorized into three types: (i) subject-object asymmetry, (ii) indicative-subjunctive asymmetry, (iii) island effects.[26]

Longobardi (1991), Bayer (1996) and Kayne (1998) noted that the there is subject-object asymmetry with regard to the possible scope of *only* in subjunctive clauses, as shown in the following examples:

(38) a. 「John has requested (that) 「Bill study only physics」」.[27]

 b. John has requested (that) 「only Bill study it」.

(38a) is ambiguous, depending on the scope of *only*. In the wide-scope reading, it means that the only request of John is that Bill study physics. In the narrow-scope reading, it means that the content of the request is that Bill study only physics. The two possible scopes are marked by the

---

[25] Multiple-syntactic-dependency with one syntactic head is not unique to FAs such as *only* and *also*. It is also available with wh-movement, which is overtly realized in certain languages.

[26] The bulk of evidence here comes from English, which has rich covert movement effects. Languages with less robust covert movement effects are discussed below.

[27] Many informants I consulted cannot get the wide scope reading with the complementizer *that* present, however. I will discuss this in chapter 4.

corner symbols.[28] On the other hand, (38b) is not ambiguous, and only has the narrow scope reading.[29] The exact nature of this subject-object asymmetry is not clear, but it is very similar to the subject-object asymmetry in classic ECP cases, and suggests that [*only* DP] undergoes covert movement.

The indicative-subjunctive asymmetry is briefly discussed in Bayer (1999).[30] It can be seen in the following examples:

(39) a. ⌜The GDR education ministry demanded that ⌜the students learn only <u>Russian</u>⌝⌝.

　　b. The GDR education ministry demands that ⌜Michael learns only <u>Russian</u>⌝.

In (39a), both the wide-scope and narrow-scope reading are available. In (39b), however, only the narrow scope reading is available. I have nothing useful to say about this asymmetry, but take this to suggest the presence of covert movement, since overt movements are known for similar asymmetries (see Szabolcsi 2006).

If our discussion of the c-command condition above is on the right track, many, if not all, of the island effects discussed in the literature can be ruled out as violations of the c-command condition. I will therefore not treat the relevant examples as arguments for covert movement.[31]

In general, the ECP and island effects, although well-known, only to some extent indicates the presence of syntactic dependency between an [FA DP/PP] and its scope-taking position, because the judgments are not always clear. Fortunately, there are other, more solid pieces of evidence for syntactic dependencies between an FA and its scope-taking position.

Locality requirement in bi-clausal sentences

When an FA is attached to a verbal projection, its scope options are even more constrained than when it is attached to DP/PP. More specifically, the scope of such an FA is limited to the

---

[28] A practice that is also used in Wagner (2006).

[29] However, according to Kayne (1998: 144), some speakers allow the wide-scope reading. He offers no explanation for this divergence. I will also leave it open.

[30] This asymmetry is obviously not shared by all speakers. Since, according to Taglicht (1984: 150), the sentence *I knew he had learnt only Spanish* can either has the wide-scope or narrow scope reading. The nature of this divergence is unclear to me.

[31] In fact, the focus of *even* is known to violate island constraints in NPI environments:

(i) They hired no linguists who had even read *Syntactic Structures*.　　　　　　　　(Rullmann 1997)

The most plausible reading is one where *even* takes scope over the entire sentence. This is problematic for theories that treat islands as barriers for movement, but not problematic for our c-command condition mentioned above. In our approach, the focus of *even* will be *Syntactic Structures*, not the whole complex NP. Since the c-command condition is not violated, they sentence should be well-formed. For more details of the special syntax of *even*, see Karttunen and Peters (1979), Rooth (1985), Herburger (2000), and Guerzoni (2003).

minimal clause that contains the FA, at least in certain varieties of English. The following examples from Taglicht (1984: 150) illustrate this point:

(40) a. I knew ⌜he had only learnt <u>Spanish</u>⌝. (I knew he hadn't learnt any other language.)[32]

    b. ⌜They were only advised to learn <u>Spanish</u>⌝. (They were not advised to learn any other language.)

In (40a), according to Taglicht, *only* can take scope over the embedded clause, but not the matrix clause. In (40b), on the other hand, *only* takes scope over the matrix clause. This suggests some sort of 'clause-mate condition' for the FA and its scope position in these cases. This again shows there is some kind of syntactic dependency between the FA and its scope position.[33]

Locality in mono-clausal sentences with multiple verbal heads

    When an FA is attached to a clausal projection, it generally cannot be separated from its scope position by more than one verbal head. To my knowledge, this generalization hasn't been previously noted in the literature.[34] Below are some examples:

(41) a. ⌜John **only** could have been dating <u>Mary</u>⌝.(He couldn't have been dating others.)[35]

    b. ⌜John could **only** have been dating <u>Mary</u>⌝. (ditto)

    c. %⌜John could have **only** been dating <u>Mary</u>⌝. (ditto)[36]

    d. %⌜John could have been **only** dating <u>Mary</u>⌝.(ditto)

(42) a. ⌜You **only** have to <u>believe it</u>⌝ if you wish to achieve it.

    b. ⌜You have **only** to <u>believe it</u>⌝ if you wish to achieve it.

    c. ?*⌜You have to **only** <u>believe it</u>⌝ if you wish to achieve it.

(43) a. ⌜**zhi**   xu   qu zuo⌝, shenghuo jiu hui gaibian   (Chinese)

       only need go act   life   JIU will change

       'You need only to act to change your life.'

    b. *⌜xu **zhi**   qu zuo⌝, shenhuo jiu hui gaibian

---

[32] There are, however, speakers who accept the matrix-clause scope reading. I will return to this in ch 4.

[33] The discussion here may be somewhat oversimplified since there are some apparent exceptions. First, it incorrectly rules out cases of Neg-raising, where the scope of the negation does not contain negation at least at overt syntax. Second, as has been discussed in note 31, *even*'s scope doesn't need to be clause-bound in certain circumstances. I will leave these issues for future research.

[34] It is certainly well-known that sentence adverbs are subject to this generalization in English (Jackendoff 1972: 76), but it seems no such observations have been extended to typical FAs like *only*.

[35] There is actually a subtle semantic/pragmatic difference between (41a) and all the other sentences (Mark Aronoff, pc). This seems to suggest that (41a) involves topicalization of *only* (see also Ernst 2002: 397).

[36] Just as there are variations of judgments among speakers about (40), there are variations of judgments here. I will return to these variations in chapter 4.

According to the scope and focus representations in (41a), it means Mary is the only one John could have been dating. The FA *only* takes wide scope over the epistemic modal auxiliary *could*. In this sentence *only* precedes *could*. In (41b), with the same focus and scope, the FA *only* occurs after the narrow-scope modal auxiliary *could* and the sentence is still fine. In (41c,d) however, two or more verbal heads intervene between *only* and its scope position, and the sentences are unacceptable with the given interpretation at least to some speakers. Examples in (42) show the same pattern with the complex modal *have to*. (43) shows that in languages like Chinese an FA cannot even be intervened by one verbal head. This again shows there is some kind of syntactic dependency between the FA and its scope position.

Overt movements in non-cleft and non-inversion sentences[37]

Overt movements show that the syntactic dependency at issue is overtly realized in some languages, and also attest to the existence of a general constraint on the availability of covert movement in a given language. This kind of movement can be found in Chinese, German, and Russian, but not in English.[38] More specifically, in Chinese-type languages, an unmoved or locally moved [FA DP] (or [DP FA]) constituent has to be interpreted in their minimal clause, the wide scope reading is never available (44a,b, 45a, 46a,b).[39] To get the wide scope reading, one can, in addition to attaching the FA to the matrix VP, move the [FA DP] constituent our of the embedded clause (44c,d, 45b,c,d, 46c,d). English does not have similar overt movements (47).

(44) a. *lisi yaoqiu xuesheng yanjou [**zhiyou** <u>yazhou de yuyan</u>] (Chinese)

      L. request student   study   only Asia   DE language

   b. lisi yaoqiu ⌜xuesheng [**zhiyou**   <u>yazhou de yuyan</u>]$_i$ cai yanjou t$_i$⌝[40]

      L. request student   only   Asia   DE language CAI study

      'Lisi requested that students only study Asian languages.'

   c. ⌜lisi [**zhiyou** <u>yazhou de yuyan</u>]$_i$ cai yaoqiu xuesheng yanjou t$_i$⌝

      L. only   Asia   DE language CAI request student   study

      'It is only <u>Asian languages</u> Lisi requested that students study.'

---

[37] Although focus-inversion and cleft-construction are available opinions in English, they do not carry the same function as focus movements in Chinese, German, and Russian. Focus-inversion is only used in formal registers, whereas focus-movements are not, and cleft-constructions are also available in languages that allow focus-movement, so they presumably carry different functions.

[38] Japanese and Korean also behave differently from English in that do not allow wide scope reading of [FA DP] in bi-clausal sentences. However, they also do not pattern like Chinese and German in that both short-distance and long-distance scrambling in the former does not have scope-related effects (cf. Futagi 2004, Lee 2004).

[39] Jacobs (1983) and Büring and Hartmann (2001) treat focusing adverbs in German as adjuncts that only attach to clausal constituents, even in cases like (45). However, as Reis (2005) has shown convincingly, there are numerous serious problems with this approach. I will hence stay with the classic approach that allows adjunction to DP.

[40] *zhiyou*, in contrast to *zhi*, is generally adjoined to non-clausal constituents.

d. 「[**zhiyou** <u>yazhou de yuyan</u>]$_i$ lisi cai yaoqiu xuesheng yanjou t$_i$ 」

       only   Asia   DE language L. CAI request student   study

     'It is only <u>Asian languages</u> Lisi requested that students study.'

(45) a. Die Studenten in der DDR wurden gezwungen 「**nur** <u>Russisch</u> zu lernen」   (German)

     the students   in the GDR were   required   only Russian to learn

     'The students in the GDR were required to only learn <u>Russian</u>.'

  b. %「Die Studenten in der DDR wurden [**nur** <u>Russisch</u>] gezwungen t$_i$ zu lernen」[41]

     the students in the GDR were   only Russian required   to learn

     'It is only Russian the students in the GDR were required to learn.'

  c. 「[**Nur** <u>Russisch</u>] wurden die Studenten in der DDR gezwungen t$_i$ zu lernen」

     only Russian were   the students   in the GDR required   to learn

     'It is only Russian the students in the GDR were required to learn.'

  d. %...weil 「die Studenten in der DDR [**nur** <u>Russisch</u>] gezwungen wurden t$_i$ zu lernen」

     since the students in the GDR only Russian required   were   to learn

     '…Since it is only Russian the students in the GDR were required to learn.'

(46) a. Ivan poprosil studentov 「izučat' [**tol'ko** <u>sintaksis</u>] 」   (Russian)

     I.   asked   students   study   only   syntax

     'Ivan asked the students to only study <u>syntax</u>.'

  b. Ivan poprosil studentov 「[**tol'ko** <u>sintaksis</u>]$_i$ izučat' t$_i$ 」

     I.   asked   students   only   syntax   study

     'Ivan asked the students to only study <u>syntax</u>.'

  c. 「Ivan [**tol'ko** <u>sintaksis</u>]$_i$ poprosil studentov 「izučat' t$_i$ 」」

     I.   only   syntax   asked   students   study

     'Ivan asked the students to only study <u>syntax</u>.' or:

     'Ivan only asked students to study <u>syntax</u>.

  d. 「[**tol'ko** <u>sintaksis</u>]$_i$ Ivan poprosil studentov 「izučat' t$_i$ 」」

     only   syntax   I.   asked   students   study

     'Ivan asked the students to only study <u>syntax</u>.' or:

     'Ivan only asked students to study <u>syntax</u>.'

(47) a. 「John forced students to 「learn [**only** English] 」」.

  b. * John forced students to [**only** English] learn.

  c. ??[**Only** English] John forced students to learn.

  d. *John [**only** English] forced students to learn.

---

[41] There are some variations among my informants about the judgments of (45b,d). The main point here is not affected, though.

The existence of overt movement of [FA DP] in Chinese-type languages provide strong evidence that there are syntactic dependency relationships between an FA and its scope position. The most natural account for the overt movement is that the relevant syntactic dependency is overtly realized in Chinese-type languages, though the details may differ in individual languages.[42] And the lack of covert movements can be naturally accounted for if we assume the existence of the overt movement strategy blocks the availability of the covert movement strategy.[43]

To sum up §3.1.2.2, it is shown that there is solid evidence that syntactic dependency relations exist between an adverbial FSE, its host, its focus, and its scope. Any alternative analysis that denies the role of syntax in the syntactic distributions of FAs cannot account for these facts. The crucial generalizations can be summarized as follows (descriptively, the B generalizations concern an FA and its host, the C generalizations concern an FA and its focus, and the D generalizations concern an FA and its scope):

(48) **Generalization A – three components**

The syntax of a focus-sensitive expression involves its dependency relations with its focus, its host, and its scope.

(49) **Generalization B1 – Free attachment, except to TP**

The first-merge host of an FA can be all kinds of syntactic categories, except a TP.

(50) **Generalization B2 – Maximal projection host**

An FA attaches only to maximal projections.

(51) **Generalization C1 – C-command Condition**

Some FAs that are attached to clausal projections (VP, TP, etc.) have to c-command their foci at overt syntax. Generally all FAs attached to non-clausal projections have to

---

[42] Overt focus-induced movements have been proposed in the literature, which usually do not involve an overt FA. The existence of vP-level scrambling/focus movement in Chinese has been noted in Ernst and Wang (1995), Shyu (1995), Zhang (1997), Soh (1998), and Tsai (2008b), ao. German *Mittelfeld*-level focus movements are discussed in Grewendorf (2005) and sources cited there, and *Vorfeld*-level focus movements such as the one we saw in (45c) are discussed in Frey (2010) and Molnár and Winkler (2010).

[43] Suzi Wurmbrand points out to me that this 'blocking effect' is not an unfamiliar concept, as in the literature similar principles have been used to account for the syntax of quantificational elements. For some discussions along the line of economy considerations, see Bobaljik and Wurmbrand (2008), Wurmbrand (2010) and works cited there. Here I will only focus on the fact that [FA DP] constituents have this property and will simply treat the variations between languages as parametric differences, leaving detailed formal treatments for future research. See also Kayne (1998) and Huang (2003) for the alternative proposal that all movements are overt and languages differ in whether remnant movements occur or not. Motivations for this approach are basically theory-internal, however (cf. Kayne 1994).

c-command their foci at overt syntax.

(52) **Generalization C2 – Adjacency**

When an FA doesn't c-command its focus, they cannot be separated by a constituent that is not part of the focus, unless other grammatical principles intervene.

(53) **Generalization C3 – Clausemate Condition 1**

When an FA doesn't c-command its focus, the focus is in the same minimal clause as the FA before it undergoes further independently-motivated A′-movements.

(54) **Generalization C4 – One-to-many association**

An FA may have more than one focus.

(55) **Generalization D1 – ECP/Island effects**

When an FA is attached to a DP or PP, the constituent [FA DP/PP] is c-commanded by its scope position and is subject to locality constraints and displays ECP effects with respect to that position.

(56) **Generalization D2 – Clausemate Condition 2**

When an FA is attached to a clausal projection, its scope is the minimal clause that contains the FA.

(57) **Generalization D3 – Intervention Condition**

When an FA is attached to a clausal projection, it can not be intervened by more than one verbal head from its scope position.

(58) **Generalization D4 – Overt movement and blocking**

If overt QR is available or scrambling has a semantic effect in a language, an [FA DP/PP] constituent cannot occur in a position that doesn't mark its scope in overt syntax.

3.1.2.3 Consequences for the syntax of sentences with multiple FAs

The descriptive generalizations compiled above not only show that several components are involved in the syntax of focus-sensitivity, but also bring important novel consequences for the syntactic hierarchical relationships between FAs in sentences with multiple FAs. This is so because according to what has been discussed so far, it is expected that those FAs may have

focus-sensitivity-related syntactic dependency relationships. On the other hand, most previous theories of adverbial syntax have ignored this possibility, and therefore would not predict that such dependencies exist (Cinque (1999), for example, has very limited discussion of focusing adverbs, basically treating them as a separate research topic, since they are not covered by his theory). It is time to test whether the consequences hold true or not.

Suppose there are two FAs, FA_1 and FA_2. FA_1 has the wide scope. There are four logically possible hierarchical relations between these two adverbs:

(59) a. FA_1 c-commands FA_2

    b. FA_2 c-commands FA_1

    c. FA_1 and FA_2 do not c-command each other

    d. Either FA can c-command the other (interchangeable word order)

In a syntactic theory without the notion of focus-sensitivity, semantic scopal properties and abstract syntactic hierarchical properties are the chief determinants of which possibilities of (59) are realized. For example, according to Ladusaw (1979, 1988), surface structural position determines the relative scopes of adverbs (as well as negation and modals), which do not undergo LF movement. Let's call this the "isomorphic approach to adverbs". Based on this assumption, several theories have developed to explain the surface structural positions of adverbs, which were reviewed in the previous chapter.[44] The crucial predictions of these theories about the syntax of multiple-FA sentences are the same: only (59a) is possible.[45] Exceptions are ascribed to other factors, such as coercion (Ernst 2002: 370).

On the other hand, a theory that addresses issues of focus-sensitivity predicts (59a-d) are all possible as long as the generalizations (48)-(58) are observed. This is so because in such a theory the overt syntax of an FA is closely associated with its focus ingredient instead of its scope relations with other FAs. By considering two novel sets of data, I will show that this prediction is borne out, and that the generalizations are on the right track.

---

[44] There are also alternative proposals to Ladusaw's analyses which argue LF movement is possible for negation (but say nothing about adverbs). They do not deal with focus-sensitivity, however. See Boeckx (2001: 536) for some discussions.

[45] The isomorphic approach, however, is usually coupled with approaches that treat FAs attached to clausal projections and those attached to non-clausal projections as separate entities (e.g. Bayer (1996)). In the latter cases, the FA can undergo covert or overt movement along with its non-clausal constituent host. Therefore, other possibilities in (59) are allowed and it seems it can still salvage the situation to some extent. The problem is that there is no theoretical grounding for treating those two occurrences of FAs as two separate entities. And even if the 'separatist' view is correct, it still has all the problems of AdvP-in-Spec approaches reviewed in chapter 2. See §3.1.2.3.2, and §3.2.2.2 for further problems of the isomorphic approach.

3.1.2.3.1 When there is at least one [FA DP]

One set of facts generally ignored in isomorphic approaches are cases that involve adverbs which take sentential scope but are attached to non-clausal constituents. As we know, this is typical for many focusing adverbs. Now let's see how it works. In a multiple-FA sentence, if at least one FA is attached to a non-clausal constituent, such as an argument DP, then the other FA can in principle occur in several positions. More specifically, the following possibilities are predicted in a monoclausal sentence (linear orders are not represented):

(60) a. …[FA$_1$ F$_{1(+N)}$]…[FA$_2$ F$_{2(+N)}$]…

    b. …[FA$_2$ vP]…, …[FA$_1$ F$_{1(+N)}$]…

    c. …[FA$_1$ vP]…, …[FA$_2$ F$_{2(+N)}$]…

    d. …[FA$_1$ TP]…, …[FA$_2$ F$_{2(+N)}$]…

    e. …[FA$_2$ TP]…, …[FA$_1$ F$_{1(+N)}$]…

    f. …[FA$_1$[FA$_2$ F$_{2(+N)}$]$_{F_1}$]…

F$_x$ represents the focus of FA$_x$. [FA$_x$ F$_{x(+N)}$] indicates that the host of the FA is a non-clausal constituent which coincides with or includes its focus, and it can occur in either the subject or object position, except when it is subject to language-specific constraints such as Generalization D4. [FA$_x$ vP] indicates that the host of the FA is a clausal constituent. These possibilities are predicted from the above generalizations, as we will discuss shortly. The only possibility that is ruled out by our generalizations is (61):[46]

(61) …[FA$_2$[FA$_1$ F$_{1(+N)}$]$_{F_2}$]…

(61) is not possible according to Generalization A and our definition of scope and focus. Since FA$_1$, by our definition, is not within the scope of FA$_2$, the former cannot be the focus or part of the focus of the latter.

Let's first consider cases that bear on (60a), repeated in (62).

(62) …[FA$_1$ F$_{1(+N)}$]…[FA$_2$ F$_{2(+N)}$]…

(63) a. %**Only**$_2$ <u>Paul</u>$_{F_2}$ is studying **even**$_1$ <u>such a fascinating language as Mohawk</u>$_{F_1}$.

    b. %?**Even**$_2$ <u>the least poisonous snake</u>$_{F_2}$ would frighten **only**$_1$ <u>Bill</u>$_{F_1}$.

---

[46] A possible exception is the [*not even* DP] construction, where *even* takes wide scope and *not* takes narrow scope. It is not clear whether *even* is part of the focus of *not*, however. I'll leave this issue aside.

c. [What language does nobody speak?]

%?**No**body$_{F_2}$ speaks **only**$_1$ Latin$_{F_1}$.

(63a) allows a reading where *even* has the wide scope (Taglicht 1984).[47] (63b) allows a reading where *only* has the wide scope (Wagner 2009). (63c) also allows the wide-scope reading for *only*. The scope reversal is predicted under Generalization D1, since the [FA DP] constituents here do not violate ECP/island constraints with regard to its scope position.[48] In languages that have the overt scrambling strategy, such as Chinese, scope reversal is expressed overtly, as expected by Generalization D4:

(64) a. [**lian**$_1$ mohuoke zheme miren de yuyan$_{F_1}$]$_i$, ye   **zhiyou**$_2$ zhangsan$_{F_2}$ zai   yanjou t$_i$

         even Mohawk so   attractive DE language YE only Z.   Prog study

         'Only$_2$ Zhangsan$_{F_2}$ is studying even$_1$ such a fascinating language as Mohawk$_{F_1}$.

   b. ***zhiyou**$_2$ zhangsan$_{F_2}$ [**lian**$_1$ mohuoke zheme miren   de yuyan$_{F_1}$]$_i$ ye zai   yanjou t$_i$

Scope reversal or not, in (63) and (64) neither FA c-commands the other. These facts are not compatible with the isomorphic assumption that adverbs do not move and are overt realizations of operators.

In (60b), repeated in (65), the narrow-scope FA is attached to a clausal constituent, while the wide-scope FA is attached to a non-clausal constituent. In this case, we expect the [FA DP] constituent to be felicitous either in the subject or object position. They are expected since nothing in the generalizations above bars their existence. However, the facts seem to defy our expectations in certain cases. Let's first consider cases where [FA$_1$ DP] occurs in the object position.

(65) …[FA$_2$ *v*P]…, …[FA$_1$ F$_{1(+N)}$]…

(66) a. [I know that Mary often buys cars. But…]

         John buys [**even**$_1$ sports cars] **often**$_2$.

   b. [You said that John often buys novels and textbooks, but I think…]

         John buys [**only**$_1$ novels] **often**$_2$.

---

[47] These examples are only acceptable to some English speakers, however. For the other speakers, the only possible way to express the meaning is to use the passive sentence: i.e. *Even such a fascinating language as Mohawk is only being studied by Mary*.

[48] As mentioned before, the behavior of *even* is different from other FAs for unknown reasons. The point here is to show that scope reversal with regard to surface word order is in principle possible in languages like English.

c. [Who didn't John see?]

   %John did**n't**$_2$ see [**only**$_1$ Mary].[49]

(67) a. [I know that Mary buys cars often. But…]

   ??John **often**$_2$ buys [**even**$_1$ sports cars].

   b. [You said that John often buys novels and textbooks, but I think…]

   *John **often**$_2$ buys [**only**$_1$ novels].

(68) a. lisi [**lian**$_1$ paoche]$_i$ dou **chang**$_2$ mai t$_i$

   L.   even sports.car DOU often buy

   'Lisi buys [**even**$_1$ sports cars] **often**$_2$.'

   b. lisi [**zhiyou**$_1$ xiaoshuo]$_i$ **chang**$_2$ mai t$_i$

   L.   only   novel   often buy

   'Lisi buys [**only**$_1$ novels] **often**$_2$.'

I assume that *often*, as a focusing adverb, is attached to *v*P (instead of a *v*P-internal complement or specifier) in (66). This means that (66) is a case of (60b), and FA$_2$ c-commands FA$_1$. This is predicted by our generalizations, but is problematic for the isomorphic analyses. (67), however, seems to be problematic for us. No generalizations so far predict them to be ill-formed. We thus need a different solution. Fortunately, the apparently problematic examples in (67) are not problematic once we consider the basic word order facts of adverbs in English. In (67), *often* is in the preverbal position, which is probably a derived position that has some kind of scopal and information structure effects (cf. Bennett 1988, Larson 2004, Kawamura 2007).[50] It could be that the narrow scope reading of *often* is incompatible with the scopal/information structure effects, which causes the degradation of (67).[51] This can be stated as the following generalization on FAs in derived positions:

(69) **Generalization E – Surface Effect**

   An FA in a derived syntactic position has interpretational effects.

---

[49] This is sentence is somewhat degraded for some speakers, perhaps due to some weak intervention effect with covert movement. The cleft version *It is only Mary John didn't see* is universally acceptable.

[50] Chomsky (1995: 48, 329), however, argues that adjuncts do not undergo topicalization, based on the fact that 'long-distance movement' is not possible. In (i), *carefully* cannot modify the car-fixing event:

(i) a. Carefully, John told me to fix the car.

   b. John told me to fix the car carefully.

However, this restriction can be reinterpreted as special interpretational effects of adverb topicalization. See note 52.

[51] Similar observations are made in Andrews (1983) and Larson (2004) about the non-ambiguity of *John twice knocked on the door intentionally*, which only has the reading where *twice* scopes over *intentionally*.

This generalization is not surprising, since it is well-known that (at least certain types of) scrambling and topicalization have closely related semantic effects.[52] Again, under an isomorphic approach these facts cannot be straightforwardly accounted for. Next, the Chinese examples in (68) are again a case of (60b), but with overt scrambling of the object DP to a position before FA_2 (the FA_2 is not in a derived position in Chinese). It is thus a case of (59c): the two FAs do not c-command each other. This again is allowed according to our generalizations. Next, consider cases where [FA_1 DP] occurs in the subject position:

(70) a. [**Even**$_1$ John] buys cars **often**$_2$.

    b. [**Only**$_1$ John] buys cars **often**$_2$.

    c.a. [**Even**$_1$ John] **often**$_2$ buys sports cars.

    b. [**Only**$_1$ John] **often**$_2$ buys sports cars.

    d. a. [**lian**$_1$ lisi]   ye **chang**$_2$ mai che

       even L.   YE often buy car

       '[even$_1$ Lisi] buys cars often$_2$.'

     b. [**zhiyou**$_1$ lisi] **chang**$_2$ mai che

       only   L. often buy car

       '[Only$_1$ Lisi] buys cars often$_2$.'

All of these examples manifest pattern (60b) and (59c), as predicted by our generalizations.

(60c), repeated in (71), is the mirror-image of (60b), where the wide-scope FA is attached to a constituent and the narrow-scope to a non-clausal constituent. This pattern is also allowed according to our generalizations. Let's look at cases where [FA_2, DP] occurs in the subject position. This should be allowed if FA_1 is not the type that is subject to Generalization C1. Examples of this kind turn out not easy to come by, however:

(71) …[FA_1 vP]…, …[FA_2 $F_{2(+N)}$]…

(72) [At school, Bill is not a popular person.]

    %[**Only**$_2$ the teacher] is **even**$_1$ willing to talk to him.[53]

(73) [I am teaching classes at this university.]

    a. *[**Only**$_2$ female students] are **usually**$_1$ very diligent.[54]

---

[52] See Kim (1991), Saito and Fukui (1998), Ernst (2002: 419), Bošković (2004) and references cited there for discussions. Generally, Quantified DPs and adverbial adjuncts are sensitive to these effects. Weak island effects that involve cases such as *Icily, he didn't speak to the lieutenant* are also a related phenomena.

[53] This sentence is not acceptable for all native speakers in this context. However, for speakers who accept (72), (73a) and (74a) are still unacceptable.

[54] According to Lee (2004: 88), Korean counterparts of this sentence may be grammatical to at least some speakers.

b. **Usually**$_1$, [**only**$_2$ female students] are very diligent.

(74) (When there is a party…)

    a. …*[**only**$_2$ John] **often**$_1$ brings beer.

    b. …**often**$_1$ [**only**$_2$ John] brings beer.

As shown in the above examples, only (72) fits (60c), where the FA attached to the clausal constituent is *even*. With FAs such as *often* and *usually*, the only way for them to scope over the FA in the subject position is to move them to the pre-subject position, as shown in (73b) and (74b). Thus it seems our generalizations are not enough, and we also need an isomorphic theory. However, since focus-sensitivity is an unmistakable syntactic phenomenon, it behooves us to supplement our generalizations instead of undermine them. I propose we add the following generalization on interplay of FAs to our arsenal:

(75) **Generalization F – Intervention Condition 2**

    a. An FA$_1$ and its scope position cannot be intervened by an [FA$_2$ DP] if any part of [FA$_2$ DP] is the focus of FA$_1$.

    b. α intervenes between β and γ if α c-commands γ and α does not c-command β.

According to this generalization, (73ab) and (74b) are ill-formed because the FA$_2$ (*only*) is part of the focus of the FA$_1$ (*usually*, *often*) and the former intervenes between the latter and the latter's scope position, which is at the edge of TP. On the other hand, (73b) and (74b) are fine since there is no such intervention. I regard cases such as (72) as an anomaly that requires independent treatment, as the English FA *even* has freer syntactic distributions than other FAs.

    For an isomorphic approach, it appears *prima facie* that the contrasts in (73) and (74) follow naturally (if one ignores the fact that *only* is not attached to a clausal constituent) since the overt syntax c-command relations of the FAs match their scope relations in the acceptable examples, manifesting (59a). However, this approach is beset by various theoretical and empirical problems. Such a theory must allow adverbs with wide scope to be base-generated either in TP- (cf. 73b, 74b) or *v*P-adjoined position, since the following sentences are fine with adverbs taking wide scope:

(76) a. Texans **often** eat barbeque.

    b. A green-eyed dog is **usually** intelligent.

This could be due to the availability of semantically vacuous scrambling of the subject DP.

c. Every female student in my class is **usually** diligent.[55]

The immediate problem is that it allows elements with a given semantic scope to be able to base-generate in more than one syntactic position. This assumption is not compatible with theories that allow only rigid base-generated positions for adverbs, or any syntactic expression, and therefore needs justification. Furthermore, to account for adverbs in vP-adjoined positions taking wide scope, one has to assume subject DPs are reconstructed in the spec-of-vP in these cases. This assumption runs afoul of the general observation that A-movements do not reconstruct (Chomsky 1995, Lasnik 1999a, Boeckx 2001). Finally, this approach makes the wrong prediction that all examples of the format [FA$_2$ DP] FA$_1$ vP should be unacceptable. To anticipate the latter discussion of sentence adverbs as FAs, the following examples falsify this prediction:

(77) a. [**Only**$_2$ John] was **luckily**$_1$ <u>rewarded by the teacher</u>$_{F_1}$.

    b. [**zhiyou**$_2$ lisi] **jingran**$_1$ <u>bu   zhidao zhe jian shi</u>$_{F_1}$ (Chinese)

      only   L.   surprisingly Negknow this Cl matter

      'Only Lisi doesn't know about this matter. I can't believe he doesn't know it.'

Due to their various C^0-related properties presented at the end of chapter 2, it follows that sentence adverbs in (77) take wide scope, while adverbs like *only* take narrow scope. In these examples, the foci of *luckily* and *jingran* do not include the subject DP and *only/zhiyou*. While these sentences are predicted to be fine by (75), since the subject DP is not part of the focus of the sentence adverb, an isomorphic/reconstruction approach would predict that they are ill-formed, since the narrow-scope *only/zhiyou* is not c-commanded by the wide-scope sentence adverb, nor does it seem possible to have the subject DP reconstructed to spec-of-vP, unless one stipulates that [FA DP] can only be reconstructed when it is not the focus of FA$_1$. This seems *ad hoc* and only further complicates the isomorphic/reconstruction approach.[56] Generalization F, on the other hand, captures the facts quite successfully. I will also show in chapter 4 that it can be derived from general, independently-motivated syntactic principles.

    Cases of (60c) where [FA DP] occurs in the object position do not involve novel predictions, so I will skip them. We have just discussed some cases of (60d) (i.e. …[FA$_1$ TP]…, …[FA$_2$ F$_{2(+N)}$]…) in (73b) and (74b), where [FA DP] occurs in the subject position. These cases are still

---

[55] See Hinterwimmer (2006), which adopts an isomorphic approach, for some relevant discussions of this sort of examples.

[56] In fact, Hinterwimmer (2006) argues that only focal DPs can be reconstructed, which is not compatible with this stipulation.

covered by our generalizations, although it is less obviously so than the other cases. This is because according to Generalization B1, TP is not a good attachment site for focusing adverbs. However, this generalization is not stated as an across-the-board constraint, so the existence of (73b) and (74b) are not counterexamples. What seems to be the problem is conceptual: Generalization B1 states TP attachment is dispreferred, while Generalization C1 states focusing adverbs should c-command their foci. There is an inherent conflict between these two generalizations when the subject DP is part of the focus of the focusing adverb. We saw that the isomorphic approach is of no help here, and our tentative solution is to posit Generalization F, according to which the intervention condition determines whether TP-adjunction or vP-adjunction is chosen. We will come back to these issues in chapter 4.

(60e) is the mirror-image of (60d), repeated in (78), where the wide-scope FA is attached to a non-clausal constituent and the narrow-scope FA is attached to TP. This pattern is patently impossible, as shown in the following examples:

(78) …[FA$_2$ TP]…, …[FA$_1$ F$_{1(+N)}$]…
(79) [You said John and Pete often buy sports cars, but I think…]
 a. ***Often**$_2$, **only**$_1$ John buys sports cars.
 b. **Only**$_1$ John **often**$_2$ buys sports cars.

Although the ill-formedness of (79a) do not follow from the generalizations we discussed in the previous section, it does follow from Generalization E (69) discussed in this section. In (79a), the FA$_2$ *often* is in a derived position, so it follows that it cannot take narrow scope with respect to an [FA DP] that is c-commanded by it.

(60f), repeated in (80), involves cases where FA$_1$ is attached to [FA$_2$ DP]. Examples of this kind are rare, but they do exist:

(80) …[FA$_1$[FA$_2$ F$_{2(+N)}$]$_{F_1}$]…
(81) a. [**Not**$_1$ [**only**$_2$ John]] left early.
 b. Our correspondents cover [**not**$_1$ [**only**$_2$ this country]] but the whole world.

It seems the scarcity of relevant examples is due to semantic factors, not syntax.

To conclude, we've seen that when we consider sentences with at least one [FA DP] constituent, an approach assuming that adverbs are all base-generated in their scope positions misses the mark. On the other hand, a focus-sensitivity-based approach correctly predicts these patterns to exist. The few apparent counterexamples (e.g. (60e) is not attested) can be accounted for by two additional, well-motivated generalizations, which further support the focus-sensitivity

approach.

3.1.2.3.2 When at least one component of an FA is the focus of another FA

Another set of facts not considered by (and incompatible with) the isomorphic approach are cases where an FA and/or its focus is the focus of another FA. These also haven't been considered by syntactic theories of focusing adverbs, which in general haven't ventured into sentences with more than one focusing adverb. More specifically, our generalizations predict (82) to be possible and (83) to be impossible.

(82) a. …[FA$_1$ vP/TP]…, [FA$_2$]$_{F_1}$…F$_2$…
    b. …[FA$_1$ vP/TP]…, […FA$_2$ …F$_2$…]$_{F_1}$…
    c. …FA$_1$…FA$_2$ …[…F$_1$…]$_{F_2}$
    d. …FA$_1$…FA$_2$…F$_{1/2}$…
(83) a. …[FA$_2$ vP/TP]…, [FA$_1$]$_{F_2}$…F$_1$…
    b. …[FA$_2$ vP/TP]…, […FA$_1$ …F$_1$…]$_{F_2}$…

Cases in (82) are possible since, again, nothing in our generalizations disallow them. Cases in (83), on the other hand, are again barred by Generalization A, which states that an FA must have a focus, and our definition of focus and scope. By our definition, FA$_1$ is not in the scope of FA$_2$, it follows FA$_1$ cannot be part of the focus of FA$_2$. Indeed, the following sentences with the scope and focus relations specified are patently semantically ill-formed:

(84) a. *He **only**$_2$ sees neighbors [**often**$_1$]$_{F_2}$.
    b. *He **often**$_2$ [does**n't**$_1$ drinks beer]$_{F_2}$.

Now let's consider cases in (82) more closely.
    In (82a), repeated in (85), the wide-scope FA is attached to the vP, and its focus is the narrow-scope FA, which can occur either in the object position, or attach to the vP. The latter cannot occur in the subject position due to Generalization F. The existence of (82a) is borne out:

(85) …[FA$_1$ vP/TP]…, [FA$_2$]$_{F_1}$…F$_2$…
(86) a. [Most people drank water at some time during yesterday's party.][57]
        John **even**$_1$ drinks [**only**$_2$]$_{F_1}$ water.

---

[57] Example from Krifka (1992).

b. [John drinks beer all the time.]

　　When he goes to a party, he **often**$_1$ drinks [**only**$_2$]$_{F_1}$ beer.

(87) [A: zhangsan changchang <u>mai xigua</u>] 　　　　　　　　　　(Chinese)

　　　'Zhangsan often buys watermelons.'

　　B: bu. ta　**zhi(you)**$_1$　[**<u>ouer</u>**$_2$]$_{F_1}$　　<u>mai xigua</u>$_{F_2}$.

　　　no he　only　　　　sometimes　buy watermelon

　　　'No. He only$_1$ <u>buys watermelons</u>$_{F_2}$ [<u>occasionally</u>$_2$]$_{F_1}$/He　buys　watermelons$_{F_2}$　only$_1$ [<u>occasionally</u>$_2$]$_{F_1}$.'

(86) illustrates cases where FA$_2$ occurs in the object position. In (87) FA$_2$ is attached to vP. In these examples, one FA is the focus of another FA. One piece of evidence for the syntactic dependency between *zhi* and *ouer* in (87) comes from the following paraphrase possibility for (87B):

(88) bu. ta　zhi(you)$_1$　[<u>ouer</u>$_2$]$_{F_1}$　**cai**　<u>mai xigua</u>$_{F_2}$.

　　　no he　only　　　　sometimes CAI　buy watermelon

　　　'No. He only$_1$ <u>buys watermelons</u>$_{F_2}$ [<u>occasionally</u>$_2$]$_{F_1}$.'

The particle *cai* here is an agreement/concord marker that appears with an FA in Chinese. In this function, it has to follow the focus. This entails that only *ouer* can be the focus of *zhi* in (88), and hence in (87B). An isomorphic approach cannot capture this paraphrase possibility.

　　　Note that in all of these examples, FA$_1$ c-commands FA$_2$, exhibiting (59a). How about cases where this c-command relation doesn't hold? In fact, it directly follows from Generalization C1 (51) that FA$_1$ nust c-command FA$_2$ in these cases. In addition, the adjacency requirement (Generalization B1) also rules out cases where FA$_2$ c-commands FA$_1$, since the FA$_2$ will not be adjacent to its focus. Thus there is no need to invoke an isomorphic approach here. The following examples illustrate this point.

(89) a. *When he goes to a party, he [**only**$_2$]$_{F_1}$ **often**$_1$ drinks <u>beer</u>$_{F_2}$. (cf. 86b)

　　b. *ta　[**<u>ouer</u>**$_2$]$_{F_1}$　**zhi(you)**$_1$　<u>mai xigua</u>$_{F_2}$. (cf. 87B)

　　　he sometimes only　　　buy watermelon

In (82b), repeated in (90), the wide-scope FA is attached to vP, and its focus includes the narrow-scope FA as well as the latter's focus. Similar to the situations in (82a), the narrow-scope FA can attach either to the object or to vP:

(90) …[FA$_1$ vP/TP]…, […FA$_2$ …F$_2$…]$_{F_1}$…

(91) [John, who is quite notorious as a party guest, did not only behave well at yesterday's party,]

he **even**$_1$ [**only**$_2$ drank [water]$_{F_2}$]$_{F_1}$.

(92) [A: lisi shujia zuo-le xie sheme?]

'A: What did Lisi do in the summer vacation?'

B: ta **zhi**$_1$ [**ouer**$_2$ [zai jia   lian   zuqiu]$_{F_2}$]$_{F_1}$.

he only   sometimes at home practice soccer

'He only practiced soccer at home occasionally. (He didn't go swimming, etc.)'

One piece of evidence for the syntactic dependency between FA$_1$ and the [FA$_2$…F$_2$] complex is that FA$_2$ in this case does not have syntactic dependencies with FA$_1$ alone, as shown by the impossibility of the following paraphrases for (92B):

(93) *ta zhi$_1$ [ouer$_2$ **cai** [zai jia lian   zuqiu]$_{F_2}$]$_{F_1}$.

he only sometimes CAI   at home practice soccer

An isomorphic approach will not be able to account for the contrast between (88) and (93).

When we consider the issue of c-command relation, we also find FA$_1$ in these cases has to c-command FA$_2$, exhibiting (59a).

(94) a. *He **only**$_2$ [**even**$_1$ drank [water]$_{F_2}$]$_{F_1}$.[58]

b. *ta [**ouer**$_2$ **zhi**$_1$ [zai jia   lian   zuqiu]$_{F_2}$]$_{F_1}$.

he sometimes only at   home practice soccer

The ill-formedness of (94) can again be accounted for by the Generalization C1. It is also ruled out by the adjacency principle (Generalization B1). Note here that this approach can achieve what all the mainstream (isomorphic) theories of adverbial syntax aim to achieve and capture the syntax of focus-sensitivity at the same time.

---

[58] However, with proper aspect and the addition of *ever*, it seems a wide-scope *even* can follow *only*. In addition, a wide-scope *even* can follow negation.

(i) [John, who used to be quite notorious as a party guest, has really changed his behavior.]
  a. He even$_1$ [only$_2$ ever drinks water]$_{F_1}$.
  b. He [only$_2$ even$_1$ ever drinks water]$_{F_1}$.
(ii) John didn't help us yesterday. He [didn't$_2$ even$_1$ know what we were doing]$_{F_1}$.

I leave these issues for future research.

In (82c), repeated in (95), part of the focus of FA_2 is the focus of FA_1. In this situation, FA_2 is usually attached to vP, while FA_1 can either attach to vP or directly to the F_1 itself. Those possibilities are illustrated in the following examples (cf. also (66)-(68)):

(95) …FA_1…FA_2 …[…F_1…]$_{F_2}$

(96) [You said that John often buys novels and textbooks, but I think…]

    a. John [buys **only**$_1$ novels$_{F_1}$]$_{F_2}$ **often**$_2$.

    b. John **only**$_1$ [buys novels$_{F_1}$]$_{F_2}$ **often**$_2$.

(97) a. lisi [**zhiyou**$_1$ xiaoshuo$_{F_1}$]$_i$ **chang**$_2$ [mai t$_i$]$_{F_2}$ (Chinese)

      L.   only   novel   often   buy

      'Lisi buys [only$_1$ novels] often$_2$.'

    b. lisi **zhi**$_1$ **chang**$_2$ [mai xiaoshuo$_{F_1}$]$_{F_2}$

      L.   only often   buy novel

      'John only$_1$ [buys novels$_{F_1}$]$_{F_2}$ often$_2$.'

In (96a) and (97a), FA_1 is directly attached to F_1, whereas FA_2 is attached to F_2, the vP, which contains F_1. [FA_1 F_1] further undergoes scope-related movement in (97a), due to the language-specific parameter setting in Chinese, as discussed in Generalization D4 (58).[59] These examples thus provide further support for our syntactic treatment of focus-sensitivity.

    Further considerations of the c-command relation between FA_1 and FA_2 seem to lead us to a puzzle, however. Nothing in our generalizations so far tells us FA_1 has to c-command FA_2 when they are both attached to vP, but in fact the c-command relation does have to hold (exhibiting (59a)):

(98) *lisi **chang**$_2$ **zhi**$_1$ [mai xiaoshuo$_{F_1}$]$_{F_2}$ (cf. 88b)

    L.   often   onlybuy novel

This seems to entail that we need a new generalization that states scope relations reflects overt syntactic c-command relations in cases such as (96b) and (97b). Nevertheless, a more careful look at these examples shows that this need not be the case.

---

[59] Similar movements are also required in German and Korean:

(i)   a. …weil   er [nur$_1$ Äpfel] oft$_2$ kauft.   (German)

       becausehe   only apple often buy

    b. *…weil er oft$_2$ [nur$_1$ Äpfel] kauft.

(ii) a. Youngsu-nun sakwa-man$_1$ congcong$_2$ santa (Korean)

      Y.-Nom   apple-only   often   buy

    b. * Youngsu-nun congcong$_2$ sakwa-man$_1$ santa

Recall our discussions of Generalization C4 (54) show that an FA has syntactic dependency relations with its *secondary* focus. If we consider our definition of focus (6) again, then we found *often* and *chang* can indeed be considered as the secondary focus of *only* and *zhi* in (96b) and (97b). In (97b), for example, the secondary focus of *zhi* includes *chang mai xiaoshuo* 'often buys novels', because a proper interpretation of the sentence involves alternative expressions such as {*chang mai xiaoshuo* 'often buys novels', *chang mai yife* 'often buys clothes', *chang mai wenju* 'often buys stationery'…}, where one of them is exclusively chosen as the actually event that does take place. I will stay with only intuitive remarks here, and leave detailed interpretational issues aside.

If the above reasoning is correct, then (96b) and (97b) should look like these:

(99) a. John **only**$_1$ [[buys <u>novels</u>$_{F_1}$]$_{F_2}$ **often**$_2$]$_{F_1}$.

   b. lisi **zhi**$_1$ [**chang**$_2$ [mai <u>xiaoshuo</u>$_{F_1}$]$_{F_2}$]$_{F_1}$

      L. only often   buy novel

Consequently, we do not need a new condition/principle to explain why FA$_1$ has to c-command FA$_2$ in these cases, the generalizations we have discussed so far (Generalizations C1, C2) are adequate for the job.

   (82d) (i.e. …FA$_1$…FA$_2$…F$_{1/2}$…) depicts the situation where two FAs share the same focus. This is an interesting case, since our generalizations so far in fact doesn't make a solid prediction about this case. This is so because the resultant structure of this pattern will involve either FA$_1$ c-commanding FA$_2$, or vice versa. Either of them seems to violate the adjacency requirement (49), which hasn't been properly defined. I will show in section 3.2 that this pattern does exist and exhibits (59b) (FA$_2$ c-commands FA$_1$). This result will be derived naturally once we adopt the theoretical framework to be discussed in chapter 4.

   In sum, considerations of cases that involve components of one FA being the focus of another FA lead to further support for our analyses of focus-sensitivity. Both the well-formed examples and ill-formed examples are expected, and all the possibilities of (59) are attested. The same cannot be said for an isomorphic approach, which assumes adverbs occupy their scope positions.

3.1.3 Summary

   In 3.1 we provided a definition of focus and focus-sensitivity, and examined in some detail the workings of focus-sensitivity in adverbial syntax. The result is that we gathered a number of working descriptive generalizations regarding sentences with one or more focus-sensitive

adverbs. Examining different cases and generalizations, we found they are all related to the three crucial components of a focus-sensitive expression: its host, its focus, and its scope. Syntactic theories that do not take these components into account face serious empirical problems. In the next section, I will show that, based on our definitions, many sentence adverbs should also be treated as focus-sensitive adverbs, and our understanding of their syntax should therefore be adjusted accordingly.

3.2 Sentence adverbs as focusing adverbs

Given the definition of sentence adverbs provided in section 2.3 and the definition and generalizations of focus-sensitivity discussed in 3.1, we are now in a position to examine whether sentence adverbs are focusing adverbs.

3.2.1 The interpretational effect

I provided basic working definitions for focus and association with focus in the beginning of this chapter. A natural first step to examine whether sentence adverbs are focusing adverbs is to see whether the former fit the definitions, repeated below:

(100) A property F of an expression α is a focus property iff F signals that alternatives of the denotation of (parts of) α are relevant for the interpretation of α.

(101) Semantic operators whose interpretational effects depend on focus are associated with focus.

To see that sentence adverbs are associated with focus, and therefore are focusing adverbs, let's first consider an example from Krifka (2007):

(102) **Fortunately**, Bill spilled [white]$_F$ wine on the carpet.

According to Krifka, a proper understanding of (102) is as follows:

(103) Among two alternatives, BILL SPILLED RED WINE and BILL SPILLED WHITE WINE, the latter one was more fortunate.

This is what (100) and (101) mean by focus and focus-sensitivity. More specifically, the focus on *white* indicates the presence of alternatives such as BILL SPILLED RED WINE and BILL SPILLED

WHITE WINE that are relevant for the interpretation of *fortunately*. Alternatively, one can also paraphrase (102) as (104), based on our more precise definition of focus in (6a), repeated in (105):

(104) Fortunately, among two alternatives, RED and WHITE, it is the latter one that is the color of wine Bill spilled on the carpet.

(105) The *focus* of a focus-sensitive expression is the expression whose denotation's substitution by alternatives is relevant for the interpretation of the FSE.

The interpretation of this sentence cannot involve an alternative such as SPILLED and DRANK, or JOHN and BILL, since the verb and the subject are not in focus.

Considerations of other types of sentence adverbs lead to the same conclusion. Take epistemic adverbs for example:

(106) a. **Probably** John [likes]$_F$ Mary.
 b. **Probably** [John]$_F$ likes Mary.

(106a, b) can be paraphrased as follows:

(107) a. Among alternatives such as LIKING, HATING, DESPISING, NOT CARING etc., the first one is the more probable attitude John has of Mary.
 b. Among alternatives such as JOHN, PETER, JENNY, etc., the first one is the more probable person who likes Mary.

This again shows that *probably* is focus-sensitive.

In addition to the paraphrase method, we can also use contexts to help us determine the focus of a sentence adverb, and to show that the sentence adverb is indeed focus-sensitive. Consider the following conversation.

(108) A: What happened?
 B: [I saw Mary give somebody some cash. Hmm…]
 i. **Perhaps** she gave [Bill]$_F$ some cash.
 ii. #**Perhaps** [she]$_F$ gave Bill some cash.
 iii. #**Perhaps** she gave Bill [some cash]$_F$.
 iv. #**Perhaps** she [gave]$_F$ Bill some cash.

136

In (108), Speaker B reports he or she saw Mary give somebody some cash, but is not sure who the recipient was. Based on some guesswork or logical reasoning, in the next sentence Speaker B offers his or her opinion about who the recipient might be. In this situation, only the recipient can be focused, as shown above, and has an interpretational effect on *perhaps*. This shows that *perhaps* is a focus-sensitive adverb.

Adverbs that are not focus-sensitive are generally identifiable by their *not* showing property (6a). Consider temporal and manner adverbs, for example:

(109) a. John spilled [white]$_F$ wine on the carpet yesterday.

b. John spilled white wine [on the carpet]$_F$ yesterday.

(110) a. John [read]$_F$ this novel quickly.

b. John read [this novel]$_F$ quickly.

In these examples, the interpretation of *yesterday* and *quickly* are not affected by which part of the sentence is focused. They modify the same events irrespective of focus. Instead, the focus can only be associated with the covert assertion operator.

I will not delve deeper into the interpretational (and relevant prosodic) issues of focus-sensitivity of sentence adverbs, since the thesis is mainly about the syntactic issues that determine their syntactic distributions. In the next section, I will provide *syntactic* arguments that many sentence adverbs are focus-sensitive.

3.2.2 Syntactic evidence

So far, we have only glimpsed the focus-sensitivity property of certain sentence adverbs from an intuitive interpretational perspective. The syntactic perspective furnishes us with further evidence. Before we begin, let's review the core properties of sentence adverbs discussed so far:

(111) a. SAs have properties of adverbial adjuncts.

b. SAs have properties of C^0 elements.

c. SAs can attach to *v*P, and sometimes to object DP.

d. SAs sometimes cannot attach to TP.

e. SAs generally do not occur lower than the 2nd auxiliary.

f. SAs generally do not occur in the sentence final position.

g. SAs generally precede other classes of adverbs.

h. The syntactic positions of SAs are dependent on the position of focus.

(111a,b) are the defining properties of sentence adverbs discussed in chapter 2. (111c-h) are the main properties of sentence adverbs that have been acknowledged and discussed in the literature, as we have shown in chapter 1 and chapter 2. Our discussions below will be able to derive most of these properties from the focus-sensitivity property of sentence adverbs. To pinpoint the focus-sensitivity property of sentence adverbs, I will utilize the descriptive generalizations established in §3.1 above.

3.2.2.1 Generalization A (an FA has four syntactic components)

As we have shown in §3.1.2.1, the core syntactic components of an FA are its focus, the host, and its scope, and it has been demonstrated clearly that expressions like *only also*, and *even* do have these components. We can now determine whether sentence adverbs have these components.

To see if there is a focus component of sentence adverbs, we need to see whether focus determines the syntactic position of sentence adverbs. First let's consider some question-answer pairs which help to identify the focus component.[60]

(112) [zhangsan sheme shihou qu-le   nuowei?]   (Chinese)[61]
      'When did Zhangsan go to Norway?'
      a. ta **yexu**   <u>qunian</u>   qu de
      b. ??**yexu** ta <u>qunian</u> qu de
      c. *ta <u>qunian</u>   **yexu**   qu de
          he last.year perhaps go-Foc
      'Perhaps he went there <u>last year</u>.'

(113) [zhangsan qunian   zuo-le sheme?]
      'What did Zhangsan do last year?'
      a. ta   qunian **yexu**   <u>sheme dou   mei zuo</u>
      b. ??**yexu**   ta   qunian   <u>sheme dou   mei zuo</u>
      c. ??ta **yexu**   qunian   <u>sheme dou   mei zuo</u>
          he perhaps last.year   what   DOU Neg do
          'Perhaps he <u>did nothing</u> last year.'

---

[60] For some unknown reason, English doesn't seem to exhibit the contrast in (112)-(115) as strongly. However, as we will see, when we consider cases that involve subject DPs marked by FAs, similar word order patterns obtain in English.

[61] Li (2005: 138) also discuss some relevant cases in Chinese. Engdahl et al. (2004) and Engels (2005) discuss similar examples in Swedish and German, respectively.

(114) [women sheme shihou qu youyong?]

'When do we go swimming?'

a. *ruguo tianqi   hao, **jiu** women <u>mingtian</u> qu

b. ruguo tianqi   hao, women **jiu** <u>mingtian</u> qu

c. #ruguo tianqi   hao, women <u>mingtian</u> **jiu** qu[62]

   if   weather good   we   tomorrow JIU go

  If the weather is good, then we do it tomorrow.'

(115) [women mingtian yao zuo sheme?]

'What do we do tomorrow?'

a. *ruguo tianqi   hao, **jiu** women mingtian <u>qu youyong</u>

b. *ruguo tianqi   hao, women **jiu** mingtian <u>qu youyong</u>

c. ruguo tianqi   hao, women mingtian **jiu** <u>qu youyong</u>

  if   weather nice we   tomorrow JIU go swim

 'If the weather is nice, then we go swimming tomorrow.'

 (112) and (113) show that an epistemic adverb such as *yexu* 'perhaps' in Chinese occurs in different positions when its focus differs. In (112), the temporal adverb *qunian* 'last year' is the focus associated with *yexu* because in the response to the question, *qunian* is among a set of alternatives that the speaker's logical guess is about. The syntactic distribution of *yexu* matches our interpretational intuition: it has to occur before its focus, which presumably allows the former to c-command the latter. In (113), on the other hand, the focus is the NegP *sheme doumei zuo* 'not do anything', which doesn't include the temporal adverb. Here, the most felicitous position for *yexu* is between the temporal adverb and the NegP, which presumably allows *yexu* to be adjacent to and c-command its focus. Thus the contrast between (112) and (113) clearly shows the existence of the focus component of epistemic adverbs such as *yexu*. (114) and (115) show similar pattern with the connective adverb *jiu*, which we argued to be a sentence adverb in chapter 2. In (114), the temporal adverb *mingtian* 'tomorrow' is the focus associated with the connective adverb *jiu*; the relationships between the antecedent clause and the consequent clause can be paraphrased as follows:

(116) If the weather is nice (tomorrow), then among the alternative time slots we can go swimming, I choose tomorrow for us to go swimming.

In (115), the *v*P *qu youyong* is the focus of *jiu*, as the conditional sentence can be paraphrased as

---

[62] In this position, *jiu* has a different function, hence the hash sign. It expresses some positive evaluation about the temporal adverb of the sentence.

in (117):

(117) If the weather is nice (tomorrow), then among the alternative things we can do tomorrow, I choose for us to go swimming tomorrow.

As shown in the examples, the syntactic position of *jiu* is just like *yexu*, is preferably adjacent to and c-commands its focus component.

The host of a sentence adverb usually overlaps with its focus, but they sometimes diverge. This can be seen in the following examples:

(118) [zhangsan yao qu nali?]
    'Where is Zhangsan going?'
    ta **yexu** yao qu <u>nuowei</u>
    he perhaps will go Norway
    'Perhaps he will go to Norway.'
(119) [fasheng-le sheme shi?]
     'What happened?'
    <u>wode che</u> **jingran** <u>bujian</u> le!
    my car surprisingly disappear Prt
    'I can't believe my car is gone!'

In (118) the focus of *yexu* is the object DP *nuowei*, and the host is *v*P. In (119), the focus is the TP *wode che bujian*, and the host is still *v*P. Thus we can see the host need not overlap with focus.

The scope component is among the most familiar properties of sentence adverbs in the literature, and I have shown in chapter 2 that there is solid evidence that sentence adverbs have properties of C^0 elements, repeated below:

(120) a. Ability to scope over the subject of the sentence.
     b. Restricted when under the scope of a clausemate C^0 element.
     c. Selection restrictions with V^0.
     d. Restricted in embedded clauses in other contexts.
     e. Clause-linking function.
     f. Denotation focus and quantification are usually not possible.
     g. Long-distance movement is not possible.

These facts unequivocally indicate the existence of the scope component of SAs.

3.2.2.2 Problems of isomorphic approaches

Before moving on to delve into more details of the focus-sensitivity property of sentence adverbs, a popular alternative family of analyses should be recalled. It has been argued in Belletti (1990) and Cinque (1999) that adverbs in general do not undergo movement, and in sentences where sentence adverbs do not occur in the sentence-initial position, it is the subject and other materials of the sentence that have undergone movements (cf. §2.2.1.2). In the same vein, there is a large body of proposals of alternative subject and object positions in Germanic languages using the position of adverbs, including sentence adverbs as a diagnostic (Diesing 1992, Vikner 1995, Bobaljik and Jonas 1996, Alexiadou and Anagnostopoulou 1998, Svenonius 2002, etc).[63] For most of these approaches, the different word order preferences in (112)-(115) are due to the presence of some information structure/scope-related movement of the subject and the temporal adverb. In (112) and (114), it would mean that the non-focused element, subject preferably undergoes movement. In (113) and (115), it would mean that the non-focus subject and temporal adverb preferably both undergo movement. More generally, it would mean that when an expression receives information focus, the non-focused elements *must* undergo movements to proper discourse-related position that are (at least) higher than TPs.

Those approaches can thus be considered as versions of isomorphic approaches, according to which adverbs occupy their scope positions overtly, and which we have shown to be problematic in dealing with sentences with multiple FAs. But even if we ignore those problems, we find other serious challenges.

A serious problem for these approaches comes from the syntax of SAs in Chinese, which makes the conditions of 'movement' impossible to state. It has been observed that in Chinese the subject-SA is the unmarked word order, and whether a sentence adverb can occur before the subject is determined by some lexical properties of adverbs, not just by the discourse status of the subject (Zhang 2000: 51, Yuan 2002, Shu 2006, Wu 2009, ao).[64] SAs are thus classified into those that can occur sentence-initially and those that cannot, as illustrated below:

---

[63] Svenonius (2002: 234) in fact proposes a somewhat 'updated' approach that doesn't rely on AgrPs, according to which neither subjects nor adverbs move across each other. Instead, they can be base-generated in either order as T′-adjuncts, modulo the filter that "an adverb may not attach to IP with a checked +Topic feature". Conceptual problems aside, this approach still shares with the older approaches the problem that the existence of the +Topic feature is unjustified. See also note 68.
[64] Wu, based on Zhang and various sources of Chinese descriptive literature, lists five factors that determine whether a sentence adverb can occur: (i) whether it is monosyllabic or disyllabic, (ii) whether it has strong or weak subjectivity, (iii) whether it has high or low degree of colloquialism, (iv) whether the subject NP bears new or old information, and (v) the different discourse-linking function of the sentence adverbs. It is clear from (i-iii) that lexical properties of the sentence adverbs affect where it can precede the subject or not.

(121) a. (**xianran**) zhangsan (**xianran**) xihuan lisi
obviously Z. obviously like L.
'Obviously Zhangsan likes Lisi.'

b. (**bijing**) hangshiji (**bijing**) hen youming
after.all Chomsky after.all very famous
'After all, Chomsky is very famous.'

c. (**daodi**) ni (**daodi**) chi-le sheme?
DAODI you DAODI eat-Pft what
'What the hell did you eat?'

d. (**yexu**) lisi (**yexu**) mingtian likai
perhaps L. perhaps tomorrow leave
'Perhaps Lisi will leave tomorrow.'

e. (**nandao**) tade che (**nandao**) huai-le?
can.it.be his car can.it.be broken-Pft
'Can it be that his car is broken?'

(122) a. (***yiding**) zhangsan (**yiding**) xihuan lisi
surely Z. surely like L.
'Zhangsan surely likes Lisi.'

b. (***ke**) jingche (**ke**) jiu zai qianmian
KE police.car KE right at front
'A police car is right in front of you! (In case you haven't noticed.)'

c. (***jingran**) lisi (**jingran**) hui jia le!
surprisingly L. surprisingly return home Prt
'I can't believe Lisi went home!/Lisi had the impudence to go home!'

d. zhiyao ta qu, (***jiu**) wo *(**jiu**) qu
as.long.as he go JIU I JIU go
'As long as he goes, I will go.'

e. (***jianzhi**) zhe (**jianzhi**) feiyisuosi!
in.effect this in.effect unthinkable
'This is practically outrageous!'

In all of the above examples, the subject is not the focus of the sentence adverb. Yet only the adverbs in (121) can freely occur either before or after the subject. If the subject-SA order esults from movement of the subject, the contrast between (121) and (122) would imply that certain adverbs force the subject to move, whereas other adverbs do not. It is mysterious why (i) adverbs trigger movement, (ii) why only some of them do, and (iii) why they don't trigger object

movement.[65]

Another serious problem for the isomorphic approach is that when we consider the facts closely, there is simply no motivation for the subject to move to derive the subject-SA word order. Movement cannot be triggered by case, since case is not related to information structure at all, but presumably a feature associated with [T].[66] The relevant movement would also have to be distinct from typical cases of topicalization, since the latter are optional operations and apply to objects as well as subjects. This can be seen in the following examples in Chinese:

(123) [shei xihuan lisi?]   (Chinese)
    'Who likes Lisi?'
    a. <u>zhangsan</u>   xihuan lisi
    b. lisi$_i$ (a),   <u>zhangsan</u> xihuan t$_i$
       Lisi Top Z.   like
     '<u>Zhangsan</u> likes Lisi./Lisi, Zhangsan likes.'

(123b) is a typical case of topicalization, where an optional topic marker can be present. The fact that (123a) is equally well-formed shows that topicalization is optional and can apply to objects. This is an obvious contrast with (112b) and (113b), where the non-focused subject follows the adverb and the sentence is less acceptable, suggesting (112a) and (113a) do not involve topicalization operations. Sentence types that are not compatible with typical topicalization in English also lead to the same conclusion:

(124) a. Has any students read the book?
      b. *Has the book any student read?
(125) a. Has any student possibly read the book?
      b. *Has possibly any student read the book?
(126) a. For Bob to like Bill would worry Harriet.
      b. *For Bill Bob to like would worry Harriet.
(127) a. For Bob to apparently like Bill would worry Harriet.
      b. *For apparently Bob to like Bill would worry Harriet.
(128) a. Charley was scared by Violet's driving the car off the cliff.
      b. *Charley was scared by the car Violet's stupidly driving off the cliff.

---

[65] Svenonius's (2002) non-movement analysis (see note 63) doesn't succeed, either. It crucially hinges on the filter that has nothing to say about the above contrasts.
[66] Presumably case can be represented as some [uT] feature on D (Pesetsky and Torrego 2001).

(129) a. Charley was scared by Violet's stupidly driving the car off the cliff.

b. *Charley was scared by stupidly Violet's driving the car off the cliff.

A natural explanation of the contrasts in (124), (126), and (128) is that TP-adjunction is barred in these contexts in English, due to the lack of a [+topic] feature at relevant heads in these contexts. [67] As a consequence, topicalization is barred in all of these contexts. The grammaticality of (125a), (127a), and (129a) as opposed to the impossibility of object topicalization therefore suggests that the former do not involve topicalization. Furthermore, the ungrammaticality of (125b), (127b), and (129b) suggests that the SA-subject word order (at least in these examples) is derived from the SAs being topicalized, a possibility allowed discussed extensively in Ernst (2002).[68] Furthermore, the proposed movement could not be a case of an 'operation of the phonological component' (Chomsky 2001), since (112)-(115) clearly show that different word orders have different semantic effects. Finally, focus movement is also out of the question, since the facts in §3.2.2.1 show sentence adverbs preferably precede the focus, rather than the other way around.

It now seems we have some strong arguments against the isomorphic approaches to the subject-SA and SA-subject word order facts: both cross-linguistic empirical facts and current theoretical frameworks do not support these approaches.

Treating SAs as FAs leads to very different predictions. There is no need to state conditions of subject movement in (121) and (122), since SA is an FA and doesn't need to occur at its scope position. Instead, its position is determined by principles that involve its scope, focus, and its host. Second, for the same reason, there is no need to find any motivation for subject movement in (112)-(115) and (125)-(129), either. What we need to find is motivation for SAs to obligatorily merge in post-subject positions in certain semantic and syntactic contexts. These predictions will be borne out when we explore further the focus-sensitive properties of SAs.

3.2.2.3 Generalization B1 (Free attachment, except TP)

A major property of FAs is that they display resistance to TP-attachment, as stated in

---

[67] Rizzi (1997: 303) proposes that a topic between the aux and the subject in sentences such as (124a) would cause HMC violation of the T-to-C movement. (See also Haegeman (2000a) and Den Dikken (2006) for similar approaches.) This proposal does not cover facts in (126)-(129), however.

[68] To salvage the isomorphic approach, Svenonius (2002) (see note 63) proposes some novel accounts of topicalization: (i) all DPs have a [+Topic] feature in English (and Danish), (ii) adverbs cannot attach to TP with a checked [+Topic] feature, (iii) sentence adverbs can either attach to TP or CP. These three proposals are supposed to account for the contrasts in (125), (127) and (129), and still allow sentence adverbs to precede the subject in other contexts. However, this approach not only fails to account for the Chinese data mentioned above , but also fails to capture the focus-sensitive property of SAs to be discussed throughout §3.2.2.

Generalization B1 in (49). This property is also manifested in sentence adverbs.

As we have just seen in the §3.2.2.2, if SAs are treated as FAs, then the facts mentioned there might be able to be accounted for by some general principles that involve the various components of FAs. Generalization B1 is exactly such a principle. Since an FA, and therefore an SA, cannot first-merge with TP, it can only occur before the subject when it is a topic. And since a language can decide which lexical item can be a topic, just as only certain manner adverbs can be topicalized in English (Jackendoff 1972: 50 ff.), it is natural that only certain FAs and SAs can be topicalized too.[69] Nothing specifically prevents SAs and FAs to attach to *v*P, so *v*P is always an available attachment site for SAs cross-lexically and cross-linguistically. Thus, it is presumable that the SAs in (121) belong to those that can be topicalized in Chinese, and those in (122) do not. This is supported by the fact that the SAs in declarative sentences in (121) can all be (optionally) followed by a topic marker *a*. This analysis combined with Generalization B1 derives the fact that only SAs in (121) can occur in the sentence-initial position. The contrasts in (125), (127), and (129) can be derived in a similar fashion. Since [+topic] feature is not available in these contexts, an SA cannot undergo topicalization, and the SA-subject word order is not possible. On the other hand, *v*P being an available attachment site, the subject-SA word order is possible. Note the focusing adverb *either*, which can presumably be topicalized (cf. (12i)), is also subject to the same condition (see also den Dikken 2006):

(130) a. John eats **either** rice or beans.

    b. **Either** John eats rice or beans.

(131) a. Does John eat **either** rice or beans?

    b. *Does **either** John eat rice or beans?

(131b) is ill-formed because TP cannot be the first-merge position for *either*, nor can T host a [+topic] feature in this context.

Another part of Generalization B1 is that FAs may attach to DPs. This predicts that FA-subject and SA-subject word order is possible as long as the subject DP is the host. This prediction is borne out:

(132) a. Does [**either** <u>John</u>] or Bill eat rice?

    b. Who did [**probably** <u>only John</u>] see?[70]

---

[69] German SAs also display similar syntactic patterns. It has been observed recently (see especially Meinunger 2006) that certain, but not all, discourse adverbs can either occur in the initial positions in V3 structures or sentence-internally in typically V2 sentences.

[70] Similar grammatical counterparts cannot be readily constructed for (126)-(129). This is presumably due to factors other than focus-sensitivity, which I will leave aside here.

In (132), the DPs *John* and *only John* are the hosts of *either* and *probably*, respectively. According to Generalization B1, these sentences should be well-formed, and they are.[71]

We conclude that Generalization B1 governs the syntax of both focusing adverbs and sentence adverbs. This supports the view that sentence adverbs are focusing adverbs. Theories that do not treat sentence adverbs as FAs have to explain why this generalization applies to them but not other types of adverbs.

3.2.2.4 Generalization C1 (C-command Condition)

Generalization C1 states that FAs tend to c-command their foci at overt syntax. We have already seen this also holds for SAs in (112)-(115)[72]. In those examples, when the focus is not c-commanded by the SA, the sentences are ill-formed. These facts show SAs are also subject to Generalization C1.

Another strong piece of evidence that SA is subject to Generalization C1 comes from the syntactic distributions of non-topicalizable SAs and SAs that occurs in sentences that don't allow topicalization. In these cases, SAs typically can only occur following the subject due to Generalization B1, as discussed in the previous section. However, when the subject is the focus of the SA[73], the SA-subject word order becomes entirely natural and even mandatory, as shown in (132) and examples below:

(133) *Who did <u>only John</u> **probably** see?

(134) a. (**yiding**) <u>hen-duo   ren</u> (*****yiding**) xihuan lisi

      surely   very-many people surely   like   L.

      'Surely <u>many people</u> like Lisi.'

  b. (**ke**) <u>you   wu-liang jingche</u> (*****ke**) zai qianmian

     KE   YOU   five-Cl   police.car KE   at front

     '<u>Five police cars</u> are right in front of you! (In case you haven't noticed.)'

  c. (**jingran**)   <u>hen-shao ren</u>   (*****jingran**)   renshi lisi!

    surprisingly very-few people surprisingly know L.

    'I can't believe that <u>very few people</u> know Lisi!'

  d. zhiyao   ta qu, *(**jiu**) <u>hen-duo   ren</u>   (*****jiu**) hui qu

    as.long.as he go   JIU   very-manypeople   JIU willgo

    'As long as he goes, <u>many people</u> will go.'

---

[71] The facts considered here are also counterexamples to Svenonius's view that all DPs are [+topic] in English.

[72] See also the references cited in note 61.

[73] Presumably for semantic reasons, the subject DP has to be a quantified DP or has an overt focus marker.

e. (**jianzhi**) <u>meiyouren</u> (***jianzhi**) renshi lisi

   in.effect nobody   in.effect know L.

   'Practically <u>nobody</u> knows Lisi.'

(132b) in the previous section shows that when the subject DP is the focus of the SA, SA can occur before the subject. (133) shows that the subject-SA word order is in fact impossible in this situation, suggesting that Generalization C1 is enforced here (due to Generalization F, to be discussed below). These facts show that a SA is *licensed* to attach to the subject DP when the latter is its focus. The Chinese examples in (134) show the same pattern. In all of the examples, the subject DPs are the foci of the SAs, as the proper interpretations of the sentences involve alternatives to the subject DPs. Again, the subject-SA word order is generally not possible, showing that Generalization C1 is enforced here, and that a SA is *licensed* to occur in a non-typical position as a result of this generalization. The details of this licensing process will be worked out in the next chapter, but the facts clearly bear on the relevance of Generalization C1.

     We can conclude that Generalization C1 applies both to typical FAs and SAs alike. Non-focusing adverbs, on the other hand, are clearly not licensable by focus and behave differently, since they are not focus-sensitive by definition.

3.2.2.5 Generalization C2 (Adjacency)

     Generalization C2 states that an FA c-commanding its focus generally has to be adjacent to it. In section 3.1.2.2, we have seen that whenever one encounters a [XP[YP[V]]] sequence, the following patterns hold (repeated from (25)):

(135) a. [FA [<u>XP</u> [<u>YP</u> [<u>V</u>]]]]

     b. [XP [FA [<u>YP</u> [<u>V</u>]]]]

     c. *[FA [XP [<u>YP</u> [<u>V</u>]]]]

It is easy to see that this generalization also applies to SAs. Consider the following examples:

(136) a. zhangsan zuotian zai tushuguan **jingran** <u>du-le shi-ben shu</u>

     b. *zhangsan zuotian **jingran** zai tushuguan <u>du-le shi-ben shu</u>

     c. *zhangsan **jingran**   zuotian   zai tushuguan <u>du-le   shi-ben shu</u>

       Z.   surprisingly yesterday at library   read-Pft ten-Cl book

     'I can't believe that Zhangsan <u>read ten books</u> at the library yesterday. (He is usually a

      slow reader.)'

(137) a. zhangsan zuotian **jingran** <u>zai tushuguan nianshu</u>

    b. *zhangsan zuotian <u>zai tushuguan</u> **jingran** <u>nianshu</u>

    c. *zhangsan **jingran**   zuotian   <u>zai tushuguan nianshu</u>

      Z.   surprisingly yesterday at library   study

    'I can't believe that Zhangsan <u>studied at the library</u> yesterday. (I thought he'd have worked in the factory.)'

(138) a. zhangsan **jingran** <u>zai baitian</u> shuijiao

    b. *zhangsan <u>zai baitian</u> **jingran**   shuijiao

      Z.   at day   surprisingly sleep

    'I can't believe that Zhangsan sleeps during daytime. (People usually sleep at night.)'

These examples show that SAs in Chinese can in principle occur before the locative adjunct and *v*P (136a), between the locative adjunct and the temporal adjunct (137a), and before the temporal adjunct (138a). However, where they actually occur in a given context is determined by the location of the focus in the sentence. They must c-command their foci (cf. (137b) and (138b)), following Generalization C1, and no preverbal elements can separate them from their foci (cf. (136b,c) and (137c)), following Generalization C2.

    Those facts show that Generalization C2 governs the syntax of both typical FAs and SAs. Non-focusing adverbs, on the other hand, occupy fixed positions, since they don't have the focus component.

3.2.2.6 Generalization C3/C4 (Clause-mate condition and one-to-many association)

    According to Generalization C4, an FA may have more than one focus. This can be observed when an FA is attached to a clausal constituent, and its foci include both an expression it doesn't c-command and an expression it does c-command, as we have seen in (36) and (37). This generalization also holds for some connective SAs. Consider the following examples ((139) and (140) are repeated from (114) and (115), but with all the relevant foci marked this time):

(139) [women sheme shihou qu youyong? 'When do we go swimming?']

    a. *<u>ruguo tianqi   hao</u>, **jiu** women <u>mingtian</u> qu

    b. <u>ruguo tianqi   hao</u>, women **jiu** <u>mingtian</u> qu

    c. *<u>ruguo tianqi   hao</u>, women <u>mingtian</u> **jiu**   qu

      if   weather good   we   tomorrow JIU go

    If the weather is good, then we do it tomorrow.'

(140) [women mingtian yao zuo sheme? 'What do we do tomorrow?']

    a. *<u>ruguo tianqi   hao</u>, **jiu** women mingtian <u>qu youyong</u>

    b. *<u>ruguo tianqi   hao</u>, women **jiu** mingtian <u>qu youyong</u>

    c. <u>ruguo tianqi   hao</u>, women mingtian **jiu** <u>qu youyong</u>

     if   weather nice we   tomorrow JIU go swim

    'If the weather is nice, then we go swimming tomorrow.'

(141) [women jintian yao qu <u>youyong</u> ma? 'Are we <u>going swimming</u> today?']

    a. <u>zhiyou ni biaoxian hao</u>, women jintian **cai** <u>qu youyong</u>

    b. *<u>zhiyou ni biaoxian hao</u>, nimen **cai** jintian <u>qu youyong</u>

    c. *<u>zhiyou ni   biaoxian hao</u>, **cai** women jintian <u>qu youyong</u>

     only   you behave   well CAI we   today go swim

    'Only if you behave yourself do we <u>go swimming</u> today.'

(142) [ni dou <u>qi dian</u> qilai ma? 'Do you always get up at <u>7 o'clock</u>?']

    a. <u>zhiyou shangban de shihou</u>, wo **cai** <u>qi dian</u> qilai

    b. *<u>zhiyou shangban de shihou</u>, wo <u>qi dian</u> **cai**   qilai

    c. *<u>zhiyou shangban de shihou</u>, **cai** wo <u>qi dian</u>   qilai

     only go.to.work DE time   CAI I   7 o'clock get.up

    'Only when I have to go to work do I get up at <u>7 o'clock</u>.'

In Chinese, *cai* and *jiu* are connective adverbs used in conditional sentences. The first one occurs in the consequent of an *only if*-clause, the second one occurs in the consequent of an *if*-clause. That these adverbs are focus-sensitive has already been established in (114) and (115), where we saw that their syntactic positions are determined by the position of their foci. However, there is one more focus that is involved in these sentences. It is the antecedent clauses themselves. This is so because sentences containing *cai* and *jiu* in their connective usage are ill-formed without the antecedent clauses, and the interpretation of these sentences always involve the implying replacing the antecedent clauses with alternatives would bring about different results. In (139)-(142), roughly speaking, if the antecedent clauses are false, the consequent clauses are either unlikely to be or cannot be true.[74] Note also that we also expect the first focus to obey Generalization C3 (clause-mate condition). This prediction is borne out:

(143) a. ⌈<u>zhiyou tianqi   hao</u>, ta **cai** <u>hui renwei women keyi qu youyong</u>⌉

      only   weather good he CAI will think   we   may go swim

      'Only when weather is good does he think we may go swimming.'

---

[74] See also Zhang (2000: 93 ff) and Hole (2004: 102) for similar analyses of Chinese connective adverbs and their other functions such as those we saw in (88) above.

b. *<u>zhiyou tianqi hao</u>, ta hui renwei 「women **cai** <u>keyi qu youyong</u>」

    only   weather good he will think   we   CAI may go swimming

The syntax and semantics of these adverbs thus parallel the syntax of *ye* 'also' as seen in (36) and (37).

    The facts show Generalization C3 and C4 hold for certain typical FAs as well as for certain SAs, suggesting they all form a natural class syntactically.

3.2.2.7 Generalization D1 (ECP/Island effects)

    Generalization D1 states a [FA DP] constituent is c-commanded by its scope position and that the former is subject to locality condition and display ECP effects with respect to the latter. The scope of the FA does not need to be at the same clause as the FA itself. SAs also display similar behaviors.[75]

(144) a. 「John was advised to learn [**probably** only French]」.
    b. John thinks 「[**probably** only Mary] learned French」.

In (144a), the expression *probably only French* can only scope over the main clause, since *probably*, being an epistemic adverb, is not semantically qualified to be in the complement of the verb *advise* (cf. **John was advised to probably go home*). When the expression occurs in the subject position (e.g. 144b), however, its scope can only be the embedded clause (the embedded clause being selected by a different verb). The syntax of *probably* thus parallels the syntax of typical focusing adverbs such as *only*. And if *probably* is not a focusing adverb, (144a) will not be grammatical at all.

3.2.2.8 Generalization D2 (Clausemate Condition 2))

    According to Generalization D2, the scope of an FA attached to a clausal constituent must be the minimal clause that contains the FA. It is easy to see that SAs are also subject to this condition in certain varieties of English:

(145) a. John thinks 「Mary **probably** went to New York」.[76]
        (=John thinks there is a likelihood Mary went to New York.)

---

[75] SAs can only occur attach to DPs with determiners and focusing adverbs. See note 73.
[76] There are some speakers who accept the wide scope reading in this sentence. See also note 32.

b. ⌜John **probably** thinks Mary went to New York⌝.
   (=There is a likelihood John thinks Mary went to New York.)

In (145a), the SA is attached to *v*P of the subordinate clause, its scope can only be the subordinate clause. In (145b), the SA is attached to *v*P of the main clause, and its scope can only be the main clause. There are no ambiguities in terms of scope here.

Thus Generalization D2 applies to typical FAs as well as to SAs.[77]

3.2.2.9 Generalization D3 (Intervention Condition)

According to Generalization D3, when an FA is attached to a clausal constituent, the FA cannot be intervened by more than one verb head from its scope position. It is well known that sentence adverbs in certain varieties of English also have this property:[78]

(146) a. George will **probably** have read the book.
     b. George will **probably** be finishing his carrots.
     c. George has **probably** been ruined by the tornado.
     d. George is **probably** being ruined by the tornado.
     e. %George will have **probably** read the book.
     f. %George will be **probably** finishing his carrots.
     g. %George has been **probably** ruined by the tornado.
     h. %George is being **probably** ruined by the tornado.

3.2.2.10 Generalization D4 (Overt movement and blocking)

According to Generalization D4, languages are parameterized by whether an [FA DP/PP] constituent can stay in situ or not. We have seen that in languages like English, such a constituent can stay in situ and presumably adopts the option of covert movement, and that the overt movement option is enforced in languages like Chinese and German, and Russia.

When it comes to SAs, we also have some evidence that this parameter carries over to them. It is never possible for an [SA DP] constituent to stay in situ in languages like Chinese (147), (148) and German (149).

---

[77] See Law (2008) for similar observations about the Chinese focus marker *shi* and the SA *daodi*.
[78] The data are from Jackendoff (1972: 76), the judgments are from my informants . Cinque (1999: 213) also reports variations among speakers about these sentences. See also note 36.

(147) a. *lisi bei bi-zhe   xue [**yexu**   <u>zhiyou fayu</u>]

       L. BEI force-Impf.learn perhaps only   French

   b. lisi [**yexu**   zhiyou fayu]$_i$ bei bi-zhe   xue t$_i$

     L. perhaps only   French BEI force-Impf.learn

     '⌜Lisi was forced to learn [perhaps <u>only French</u>]⌝.'

(148) a. *lisi chang mai [**dagai**   zhiyou pingguo]

       L. often buy probably only   apple

   b. lisi [**dagai**   zhiyou pingguo]$_i$ chang mai t$_i$

     L. probably only   apple   often buy

     'Lisi buys [probably only apples] often.'

(149) a. *Die Studenten wurden gezwungen [**wahrscheinlich** nur <u>Russisch</u>] zu lernen

      the students   were   required   probably   only Russian learn

   b. Die Studenten wurden [**wahrscheinlich** nur <u>Russisch</u>]$_i$ gezwungen t$_i$ zu lernen

     the student   were   probably   only Russian required   learn

     'It is probably only Russian the students were required to learn.'

   c. [**Wahrscheinlich** nur <u>Russisch</u>]$_i$ wurden die Studenten gezwungen t$_i$ zu lernen

     'It is probably only Russian the students were required to learn.'

   d. weil die Studenten [**wahrscheinlich** nur <u>Russisch</u>] gezwungen wurden t$_i$ zu lernen

     '…since it is probably only Russian the students were required to learn.'

   e. ***Wahrscheinlich** die Studenten wurden [nur <u>Russisch</u>]$_i$ gezwungen t$_i$ zu lernen

     probably   the student   were   only Russian required   to learn

None of these languages allows the option in English as shown in (144a), where [SA DP] can overtly occur in a position that doesn't match its scope position. Instead, overt movement/ scrambling or other scope marking strategies are adopted. Note that although alternative analyses are possible for most cases here, where the SAs can be treated as adjuncts to clausal constituents (since they are preverbal), the fact that (149c) is grammatical strongly suggests that the SA is attached to DP, instead of to a clausal constituent, due to the V2 requirement in German (cf. 149e). This is exactly what we expect if SAs are FAs.

3.2.2.11 When there is a sentence adverb and an FA in a sentence

    As shown in §3.1.2.3, another set of facts that distinguish FAs from non-FAs are sentences that contain multiple FAs. According to our predictions, all the options in (59) are possible, repeated below:

(150) a. FA$_1$ c-commands FA$_2$

b. FA$_2$ c-commands FA$_1$

c. FA$_1$ and FA$_2$ do not c-command each other

d. Either FA can c-command the other (interchangeable word order).

Suppose now that FA$_1$, the FA that taking wide scope, is a sentence adverb, we expect the same pattern to hold. On the other hand, if sentence adverbs are not FAs, we expect only (150a) to be possible. To see if our predictions are borne out, let's re-examine two sets of data.

3.2.2.11.1 When there is at least one [FA DP]

As has been shown in §3.1.2.3.1, there are six logical possibilities when there is at least one [FA DP] constituent in sentences containing two FAs:

(151) a. …[FA$_1$ F$_{1(+N)}$]…[FA$_2$ F$_{2(+N)}$]…

b. …[FA$_2$ vP]…, …[FA$_1$ F$_{1(+N)}$]…

c. …[FA$_1$ vP]…, …[FA$_2$ F$_{2(+N)}$]…

d. …[FA$_1$ TP]…, …[FA$_2$ F$_{2(+N)}$]…

e. …[FA$_2$ TP]…, …[FA$_1$ F$_{1(+N)}$]…

f. …[FA$_1$[FA$_2$ F$_{2(+N)}$]$_{F_1}$]…

Now let's examine what happens when a sentence adverb enters the mix as a wide-scope FA. First, (151a) is attested. Consider the following sentences:

(152) [Does John speak only English?]

No. [**Probably**$_1$ nobody] speaks **only**$_2$ English.

(153) [What language does nobody speak?]

%**No**body$_2$ speaks [**probably**$_1$ only Latin].[79]

In both (152) and (153), SA *probably* takes the wide scope and is attached to a DP constituent, which already contains another FA, negation. The examples show that an [SA DP] constituent behaves just like a typical [FA DP] constituent in that they both can occur in either the subject or object position and still take the wide scope. In languages that adopt the overt movement strategy, such as Chinese, the wide-scope [SA DP] constituent has to be in a position that c-commands the

---

[79] English speakers tend to prefer the cleft *it is probably only Latin nobody speaks*, but (153) is still acceptable to some speakers.

narrow-scope [FA DP], just as expected:

(154) [sheme yuyan meiyouren shuo? 'What language does nobody speak?']

[**yexu**$_1$ zhiyou ladingwen]$_i$ **meiyou**ren$_2$ shuo   t$_i$
  perhaps only   Latin   nobody   speak
  'It is perhaps$_1$ only Latin nobody$_2$ speaks/Nobody$_2$ speaks perhaps$_1$ only Latin.'

In these examples, neither the SA nor the FA negation c-commands the other. They thus belong to case (150c).

(151b) (repeated in (155)) is also attested. For brevity, let's just consider cases where [SA DP] is in the object position:

(155) …[FA$_2$ vP]…, …[FA$_1$ F$_{1(+N)}$]…
(156) [Who didn't John see?]

%John did**n't**$_2$ see [**probably**$_1$ only Mary].
  b. [You said that John often buys novels and textbooks, but I think…]
    %John buys [**probably**$_1$ only novels] **often**$_2$.

These examples are not completely acceptable, but they have various degrees of acceptance among different speakers.[80] The reason they are to some extent degraded is perhaps due to some lexical properties of sentence adverbs and some weak intervention effect on covert movements. But since these sentences are not completely out for all speakers, I will take them to be evidence that (151b), and therefore (150b), are attested, and that SA behaves like typical FAs. Similar examples can be constructed in German (where the constituency is made clear by the V2 effect), where overt movements have to take place:

(157) a. ?[**Wahrscheinlich**$_1$ nur Peter] hat Maria **nicht**$_2$ gesehen.
    probably   only P.   has M.   not   see
    'Maria didn't$_2$ see [probably$_1$ only Peter].
  b. ?[**Wahrscheinlich**$_1$ nur Äpfel] kauft Peter **oft**$_2$
    probably   only apple buys P.   often
    'Peter buys probably$_1$ only apples often$_2$.'

Again, these examples are slightly degraded due to some weak intervention effect, just like

---

[80] Again, the cleft version is fully acceptable (i.e. *It is probably only Mary John didn't see*).

English examples. However, since they are still to some extent acceptable, I take them to be evidence for (151b).

(151c), repeated in (158), is also attested, but it is limited. When [FA$_2$ F$_{2(+N)}$] occurs in the object position, the sentences are fully acceptable, as is well known. When [FA$_2$ F$_{2(+N)}$] occurs in the subject position, however, the sentences are not always well-formed:

(158) …[FA$_1$ *v*P]…, …[FA$_2$ F$_{2(+N)}$]…
(159) a. *[**Only**$_2$ John]$_{Fı}$ has **probably**$_1$ written emails to Mary.
    b. *[**No**$_2$body]$_{Fı}$ has **unfortunately**$_1$ passed the GRE exam.

These examples are all ill-formed if the subject DP is part of the focus of the SA. As we have seen earlier, their ill-formedness is not just due to the scope factor, it is due to the interaction of scope and focus that derive their ill-formedness. This is our Generalization F (intervention effect). On the other hand, there are also examples where the subject DP is not part of the focus of the SA. These examples are well-formed, as we have seen in (77), repeated below:

(160) a. [**Only**$_2$ John] was **luckily**$_1$ <u>rewarded by the teacher</u>$_{Fı}$.
    b. [**zhiyou**$_2$ lisi] **jingran**$_1$ <u>bu zhidao zhe jian shi</u>$_{Fı}$ (Chinese)
      only   L.   surprisingly Negknow this Cl matter
      'Only Lisi doesn't know about this matter. And I can't believe that he doesn't'

These examples manifest yet another case of (150c).

Case (151d) and (151f), repeated in (161a,b), are usually not easy to distinguish when [FA$_2$ F$_{2(+N)}$] occurs in the subject position and the SA precedes it. When we look at the examples in English in (162), it seems we can either treat the SAs as TP-adjuncts or DP-adjuncts, since both positions are independently attested in English (163).

(161) a. …[FA$_1$ TP]…, …[FA$_2$ F$_{2(+N)}$]…
    b. …[FA$_1$[FA$_2$ F$_{2(+N)}$]$_{Fı}$]…
(162) a. **Probably**$_1$ [**only**$_2$ John]$_{Fı}$ has written emails to Mary.
    b. **Probably**$_1$ [**no**$_2$body]$_{Fı}$ has written emails to Mary.
    c. **Unfortunately**$_1$ [**no**$_2$body]$_{Fı}$ has passed the GRE exam.
(163) a. Mary saw [**probably** [$_{DP}$ only John.]]
    b. [**Probably** [$_{TP}$ it is raining.]]

I will hence treat SAs in (162) as either DP- or TP-adjuncts, and regard them as evidence that

(151d) and (151f) are attested. These cases all manifest pattern (150a).

Finally, (151e), repeated in (164), is not attested.

(164) …[FA$_2$ TP]…, …[FA$_1$ F$_{1(+N)}$]…

(165) a. ***Often**$_2$, [**probably**$_1$ only John] buys beer.

b. [**Probably**$_1$ only John] **often**$_2$ buys beer.

(165a) is ill-formed. As we have seen, Generalization E correctly rules out this possibility. This result therefore shows again that SAs behave like typical FAs.

To conclude, the facts examined in this section supports an analysis in which sentence adverbs are focus-sensitive expressions, since they in general exhibit the properties exhibited by typical focusing adverbs.

3.2.2.11.2 When at least one component of an FA is the focus of another FA

As we have seen in §3.1.2.3.2, there are six possibilities when at least one component of an FA is the focus of another FA. Four of them are allowed by our generalizations, and two of them are barred. Specially, the cases in (166) are allowed by our generalizations, whereas the cases in (167) are barred by our definition of scope.

(166) a. …[FA$_1$ vP/TP]…, [FA$_2$]$_{F_1}$…F$_2$…

b. …[FA$_1$ vP/TP]…, […FA$_2$ …F$_2$…]$_{F_1}$…

c. …FA$_1$…FA$_2$ …[…F$_1$…]$_{F_2}$

d. …FA$_1$…FA$_2$…F$_{1/2}$…

(167) a. …[FA$_2$ vP/TP]…, [FA$_1$]$_{F_2}$…F$_1$…

b. …[FA$_2$ vP/TP]…, […FA$_1$ …F$_1$…]$_{F_2}$…

This appears correct. Cases following the patterns in (167) are indeed unattested:

(168) a. *John **only**$_2$ [**probably**$_1$]$_{F_2}$ saw Mary.

b. *John **often**$_2$ [**actually**$_1$ brings <u>beer</u>$_{F_1}$]$_{F_2}$.

By contrast, (166a) is attested, as shown in (169):

(169) a. Mary **obviously**$_1$ likes [**only**$_2$]$_{F_1}$ Bill.

b. Mary **probably**$_1$ [**always**$_2$]$_{F_1}$ walks to school.

In (169a), SA is attached to vP, while FA_2 is attached to object DP. In (169b), FA_2 is attached to vP. Again, as expected, SAs have to c-command FA_2 in these examples due to the c-command c-condition and the adjacency requirement:

(170) a. *Mary [**always**$_2$]$_{F_1}$ **probably**$_1$ walks to school.
　　 b. *Mary [**only**$_2$]$_{F_1}$ **obviously**$_1$ likes Bill.

We also expect FA_2 not to occur in subject DP in these cases, due to Generalization F, this is again borne out:

(171) a. *[[**only**$_2$]$_{F_1}$ Mary] **obviously**$_1$ likes Bill.
　　 b. *[[**No**$_2$]$_{F_1}$ student] will **probably**$_1$ read this book.

In (166b), repeated in (172), both F_2 and FA_2 are the focus of FA_1. Cases of this kind are very similar to what we have shown above with (166a), with the slight modification that F_2 is now also included as part of F_1 (the context will have to be modified accordingly):

(172) …[FA_1 vP/TP]…, […FA_2 …F_2…]$_{F_1}$…
(173) a. Mary **obviously**$_1$ likes [**only**$_2$ <u>Bill</u>$_{F_2}$]$_{F_1}$.
　　 b. Mary **probably**$_1$ [**always**$_2$ <u>walks to school</u>$_{F_2}$]$_{F_1}$.
　　 c. They certainly learnt Spanish and they **probably**$_1$ [**also**$_2$ learnt <u>French</u>$_{F_2}$]$_{F_1}$.
　　 d. I had hoped either to speak to Bill myself, or if he was out to leave a message with his wife; but **unfortunately**$_1$ [his wife was **also**$_2$ <u>out</u>$_{F_2}$]$_{F_1}$.[81]

In (166c), repeated in (174), part of the focus of FA_2 is the focus of FA_1. In this situation, FA_2 is usually attached to vP, while FA_1 can either attach to vP or directly to the F_1 itself. The following examples illustrate relevant cases:

(174) …FA_1…FA_2 …[…F_1…]$_{F_2}$
(175) lisi chang mai sheme?
　　 [What does Lisi buy often?]
　　 a. lisi **yexu**$_1$　　**chang**$_2$ [mai <u>xiaoshuo</u>$_{F_1}$]$_{F_2}$
　　　 L.　perhaps　often　buy novel
　　　 'Lisi perhaps buys novels often.'

---

[81] (173c,d) are taken from Taglicht (1984: 162-163).

b. ?lisi [**yexu**$_1$ xiaoshuo$_{F_1}$]$_i$ **chang**$_2$ [mai t$_i$]

   L. perhaps novel   often   buy

'Lisi buys [perhaps novels] often.'[82]

In (175a), both SA and FA$_2$ are attached to vP. In (175b), SA is attached to DP, while FA$_2$ is attached to vP. Again the situation is just like what we observed with typical FAs above. Note here that SA cannot be c-commanded by FA$_2$:

(176) *lisi **chang**$_2$ **yexu**$_1$ [mai xiaoshuo$_{F_1}$]$_{F_2}$

   L. often   perhaps   buy   novel

As we discussed earlier, the ill-formedness of (176) follows from Generalization C4 (one-to-many association). Here we have the same situation. *chang* 'often' here is in fact part of the secondary focus of *yexu* 'perhaps', the former therefore has to be c-commanded by the latter when both are attached to vP. This solution predicts when FA$_2$ is not part of the focus of the SA, the FA$_2$ > SA order is possible. We will see soon this prediction is borne out.

     Finally, let's consider (166d), repeated in (177). In this case, the SA and FA$_2$ share the same focus. Examples of this kind are in fact commonplace, although they haven't been discussed in the literature to my knowledge. This is mainly due to the fact linguists have paid little attention to the focusing property of various common adverbs. Consider the following sentences in Chinese:[83]

(177) …FA$_1$…FA$_2$…F$_{1/2}$…

---

[82] The English translation is again not so good perhaps due to some weak intervention effect.

[83] Here I only list examples in Chinese. Although English exhibits similar patterns, English seems to extraordinarily allow much more freedom in terms of adverb ordering:

 (i) a. We are **still**$_2$ **probably**$_1$ north of Princeton. (Ernst 2002: 370)

     b. We are **probably**$_1$ **still**$_2$ north of Princeton.

 (ii)a. Pollution will **always**$_2$ **probably**$_1$ exist.

     b. Pollution will **probably**$_1$ **always**$_2$ exist.

(iii)a. He'd **never**$_2$ **probably**$_1$ have enough courage to leave.

     b. He'd **probably**$_1$ **never**$_2$ have enough courage to leave.

(iv)a. I **only**$_2$ **really**$_1$ dance sitting down.

     b. I **really**$_1$ **only**$_2$ dance sitting down.

In Chinese counterparts of (i)-(iv), only SA > FA$_2$ is the acceptable order. I believe the possibility of right-adjoining an FA to its focus in English is the reason for the difference between the two languages. Again, I will leave the details open here.

(178) a. ta **shuobuding**₁ **shenzhi**₂ [qu-guo xinjiapuo]$_{F_1/F_2}$

    b. ta **shenzhi**₂ **shuobuding**₁ [quguo xinjiapuo]$_{F_1/F_2}$

    he even   maybe   go-Exp Singapore

    'Maybe he has even been to Singapore.'

(179) [Lisi doesn't know how to sharpen a pencil…]

    a. ta **jingran**₁   **changchang**₂   [yong ya yao   qianbi]$_{F_1/F_2}$

    b. ta **changchang**₂ **jingran**₁   [yong ya   yao   qianbi]$_{F_1/F_2}$

    he often   surprisingly use tooth bite   pencil

    'He often uses his teeth to bite pencils. I can't believe someone can do such a thing.'

(180) [His idea is stupid, but…]

    a. dajia   **jingran**₁ **dou**₂   [zancheng tade   yijian]$_{F_1/F_2}$

    b. dajia   **dou**₂   **jingran**₁   [zancheng tade   yijian]$_{F_1/F_2}$

    everyone DOU   surprisingly agree.with his   idea

    'Everyone agrees with his idea. I can't believe someone can agree with that!'

In these sentences, SA and FA₂ can freely occur either in SA > FA₂ or FA₂ > SA order.[84] Facts like these lead Ernst (2002: 371) to propose that propositions expressed by the [Spkr-Or + sentence] sequence can be *coerced* into events.[85] However, this account faces robust systematic counterexamples. First, when FA₂ is *only* or *not*, the FA₂ > SA order becomes impossible:[86]

---

[84] See also H. Huang (1990) and Yuan (2002) for some more examples of this kind in Chinese.

[85] Ernst in fact discusses only (178) and some English examples, treating *shenzhi* as FA₁. It is nevertheless clear that the data discussed here would also have to receive a coercion treatment in his analysis due to scope considerations. The limitations of this approach in fact lead Ernst to abandon it in Ernst (2009), adopting instead a semantic treatment of sentence adverbs as positive polarity items. This analysis however has even more serious problems, since it fails to cover data that involve focusing adverbs that are not negative adverbs, and it also fails to account for the facts in note 83.

[86] Certain SAs that are suffixed by *-de*, however, behave differently:

(i) zhanzhen hou, lisi **zhi xingyunde** goucun-le   yi-tiao xiaoming

   war   after L. only luckily   preserve-Pft one-Cl little.life

   'After the war, Lisi only had the luck to stay alive (he didn't have any other things left).'

(ii) lisi **meiyou buxingde**   shoushang

   L. Neg   unluckily get.hurt

   'Lisi didn't have the bad luck to get hurt.'

As shown in the translation, the semantics (and probably syntax) of these adverbs seems to be different from the other SAs, which may lead to their peculiar syntactic distributions. I leave this for future research.

(181) a. ta **shuobuding**$_1$ **zhi**$_2$ qu-guo xinjiapuo
     b. * ta **zhi**$_2$ **shuobuding**$_1$ qu-guo xinjiapuo
        he only maybe   go-Exp Singapore
        'He has maybe only been to Singapore.'

(182) a. ta **jingran**$_1$ **meiyou**$_2$ yong ya   yao qianbi
     b. *ta **meiyou**$_2$ **jingran**$_1$ yong ya   yao qianbi
        he Neg   surprisingly use tooth bite pencil
        'I can't believe he didn't use his teeth to bite pencils!'

Second, when the FA$_2$ is a frequency or quantificational adverb and the SA is an epistemic
adverb, the FA$_2$ > SA order is not possible.

(183) a. ta **yexu**$_1$ **changchang**$_2$ yong ya   yao   qianbi
     b. *ta **changchang**$_2$ **yexu**$_1$ yong ya   yao   qianbi
        he often   perhaps use tooth bite   pencil
        'Perhaps he often use his teeth to bite pencils.'

(184) a. dajia   **yexu**$_1$ **dou**$_2$ zancheng   tade   yijian
     b. *dajia   **dou**$_2$ **yexu**$_1$ zancheng   tade   yijian
        everyone DOU   perhaps agree.with   his   idea
        'Perhaps everyone agrees with his idea.'

Third, when contexts indicate FA$_2$ is part of the focus of SA, then only SA > FA$_2$ order is
possible.

(185) [wo yiwei ta zonglai bu   chumen… 'I thought he never went out…']
     a. ta **jingran**$_1$ [**changchang**$_2$ qu xianggang]$_{F_1}$
     b. *ta [**changchang**$_2$ **jingran**$_1$ qu xianggang]$_{F_1}$
        he often   surprisingly go Hong.Kong
        'I can't believe (as it turns out) he goes to Hong Kong often.'

Ernst's coercion approach cannot capture the systematic constraints against FA$_2$ > SA order in
(181)-(185). If a proposition expressing [Spkr-Or + sentence] sequence can be *coerced* into an
event, it is unclear what prevents the new 'events' from being modified by *only* or *not*, and what
makes frequency adverbs and epistemic adverbs problematic for the coercion process, and what
makes context affect the coercion potential. However, if we follow the main proposal that
sentence adverbs are focusing adverbs and have syntactic dependency relations with their foci,

160

all the above facts fall into place.

I propose that in (178b), (179b), and (180b), the *v*Ps are in fact the foci of both the SA and FA$_2$. In all of these sentences, the *v*P constituent is not only the focus of the aspectual, distributive, or frequency focusing adverb, it is also the focus of the SA. More specifically, in (178b), repeated in (186), the speaker is expressing the following things:

(186) ta **shenzhi**$_2$ **shuobuding**$_1$ [quguo xinjiapuo]$_{F_1/F_2}$

(187) a. yiban   wo   buhui   xiangdao$_2$ ta [qu-guo   xinjiapuo]$_{F_2}$

     generally   I   Neg.Subj expect   he go-Exp   Singapore

     'Generally I wouldn't expect he has been to Singapore.'

   b. ta shuobuding$_1$ [qu-guo   xinjiapuo]$_{F_1}$

     he maybe   go-Exp   Singapore

     '(With new evidence) He has maybe been to Singapore.'

In other words, the part of the proposition *qu-guo xinjiapuo* 'has been to Singapore' is the main focus of the epistemic adverb and the evaluative adverb *even*. Crucially, there is no need for *shenzhi* 'even' to be part of the focus of *shuobuding* 'maybe'. This is so because evaluation is not an inherent part of the event for the purpose of deciding whether a sentence is true or not. Similarly, in (179b), the speaker is making the following points:

(188) ta **changchang**$_2$ **jingran**$_1$ [yong ya   yao   qianbi]$_{F_1/F_2}$

(189) a. ta   changchang$_2$ [yong ya   yao   qianbi]$_{F_2}$

     he   often   use   tooth bite   pencil

     'He often uses his teeth to bite pencils.'

   b. youren   jingran$_1$ hui [yong ya   yao qianbi]$_{F_1}$

     one   surprisingly Subj. use   tooth bite pencil

     'I can't believe someone would use his teeth to bite pencils.'

According to (189), the speaker of (179b) expresses the proposition that Lisi often uses his teeth to bite pencils, and a part of the proposition *yong ya yao qianbi* 'use his teeth to bite pencils' is surprising. Here, there is no need for *changchang* 'often' to be part of the focus of *jingran* 'surprisingly', because frequency is not necessarily relevant for the purpose of deciding whether the nature of the event itself is surprising or not.

With these basic semantic anatomies of relevant examples at hand, we are now in a position to explain the facts in (181)-(185).

In (181) and (182), repeated in (190) and (192), the FA$_2$ is *zhi* 'only' and *meiyou* 'not',

161

respectively. For SA and FA$_2$ to have the same focus in these sentences, we would have to paraphrase (181b) and (182b) as (191) and (193), respectively:

(190) * ta **zhi$_2$ shuobuding$_1$** [qu-guo xinjiapuo]$_{F_1/F_2}$

(191) a. ta zhi$_2$ [qu-guo xinjiapuo]$_{F_2}$

    he only go-Exp Singapore

    'He has only been to Singapore.'

    b. ta shuobuding$_1$ [qu-guo xinjiapuo] $_{F_1}$

     he maybe   go-Exp Singapore

     'He has maybe been to Singapore.'

(192) *ta **meiyou$_2$ jingran$_1$**   [yong ya   yao qianbi]$_{F_1/F_2}$

(193) a. ta   meiyou$_2$ [yong ya   yao qianbi]$_{F_2}$

    he Neg   use tooth bite pencil

    'He didn't use his teeth to bite pencils.'

    b. ta jingran$_1$ hui [yong ya   yao qianbi] $_{F_1}$

     he surprisingly Subj. use tooth bite pencil

     'I can't believe he would use his teeth to bite pencils.'

Clearly, the paraphrases show that the sentences are ill-formed for semantic/pragmatic reasons. One cannot assert the exhausitivity of an event (191a) when making an epistemic statement about the same event (191b) at the same time, because this would involve simultaneous assertion and epistemic qualification. Similarly, one cannot assert the negation of an event (193a) while expressing an evaluation of the same event (193b) at the same time. This would involve simultaneous assertion of negative truth-value and presupposition of positive truth-value. In other words, (181) and (182) are out because they induce semantic/pragmatic incoherence.

    In (183), repeated in (194), FA$_2$ is a frequency adverb and the SA is an epistemic adverb. For both adverbs to have the same focus, we would have to paraphrase (183b) as follows:

(194) *ta **changchang$_2$ yexu$_1$**   [yong ya   yao   qianbi]$_{F_1/F_2}$

(195) a. ta yexu$_1$ changchang$_2$ [yong ya   yao qianbi]$_{F_2}$

    he perhaps often   use tooth tooth pencil

    'Perhaps he often use his teeth to bite pencils.'

    b. ta yexu$_1$ [yong ya   yao qianbi]$_{F_1}$

     he perhaps use tooth bite pencil

     'Perhaps he uses teeth to bite pencils.'

Again, the paraphrase shows that the sentence is ill-formed for semantic reasons. One cannot assert the frequency of an event type when making an epistemic statement about a specific token of the event.

Finally, we can see the role of focus is clearly relevant in (185). The context makes it clear that *changchang* 'often' is part of the proposition that leads to the speaker's surprise, therefore it is necessarily part of focus of *jingran* 'surprise'. Not focusing on *changchang* would make the sentence semantically/pragmatically anomalous.

If the above reasoning is on the right track, then we have solid evidence for the existence of (166d), where SA and FA_2 share the same focus. This in turn is a strong piece of evidence that sentence adverbs are focusing adverbs. If they are not focusing adverbs, the $FA_2 >$ SA order will need to be accounted for by some mechanism like coercion, which cannot capture the facts above that are systematically related to focus. Note in the cases discusses here, there is a preferred order $FA_2 >$ SA when they both share the same focus. So far, our generalizations fail to predict this preference. We thus have a new generalization on ordering of FAs to be accounted for:

(196) **Generalization G – Late insertion effect**

When two FAs share the same focus, the FA with narrow scope c-commands the FA with wide scope.

I will discuss this generalization in more detail in the next chapter. Its presence can be derived naturally once we consider the overall architecture of grammar.

We have thus gone over all the major logical possibilities we may have with regard to the syntactic relations between two FAs, one of them being a sentence adverb. The facts again all indicate that sentence adverbs are focusing adverbs.

3.3 Conclusion

In this chapter we examined a different set of syntactic properties of sentence adverbs. First we see that typical focusing adverbs have a set of syntactic properties that involves dependency relations between their focus, their host, and their scope. Then we see that when two focusing adverbs interact, their syntactic distributions conform to these dependency relations. In addition, we also examined cases where the presence of an FA 'intervenes' between certain dependency relationships. The result of all these is a set of descriptive generalizations about the syntax of these dependency relationships. Treating these descriptive generalizations about FAs as diagnostics, we went on to examine whether sentence adverbs have these properties. With careful scrutiny of their semantic and syntactic properties, we found they do match the basic definitions

of focusing adverbs and the descriptive generalizations associated with focus-sensitivity. These findings consolidate our understanding of sentence adverbs. Crucially, they also show that any syntactic account based on scope positions alone cannot be correct since such an analysis has nothing to say about focus-sensitivity at all. In the next chapter, we will see how these properties, as well the properties addressed in previous chapters, can be naturally accounted for in the current minimalist theoretical framework.

4. An Agree Analysis of Sentence Adverbs

The previous chapters have reviewed the major properties of sentence adverbs. Chapter 2 established that the term 'sentence adverb' denotes a theoretically important set of expressions. Chapter 3 established that these expressions are focus-sensitive. With theoretical footing and factual basis established, we now have our core explicanda. The next step is to provide a theoretical account of all these properties if possible. To recap, we need to account for the following facts:

(1) a. The syntax-semantics mismatch problem.
      b. The theoretical status of adverbial adjuncts.
      c. The C^0 properties of sentence adverbs.
      d. The syntax of focusing adverbs.
      e. Sentence adverbs are a heterogeneous group.
      f. Cross-linguistic variation.

In this chapter I offer an Agree analysis of sentence adverbs and relevant phenomena. The main proposal is that sentence adverbs, as well as focusing adverbs, are inflectional affixes writ large. Their presence is derived from two independent syntactic expressions having a connection between each other in terms of their feature makeup. One expression is a covert C^0 element, which hosts the *probe*, an interpretable [iMood] feature. The other expression is a T^0, v^0, or an X^0 element that hosts the *goal*, an uninterpretable and unvalued [uMood] feature. The Agree operation between the probe and goal triggers *delayed-Merge*, realized as late insertion of an SA at the edge of the head bearing *goal* or a projection of the head (the latter being an instance of *Pied-pipe*). In addition, movement of the phrase adjoined to the SA may also take place as a case of typical A′-movement. Since the host of the 'affix' can be a phrase, and the 'affix' itself can be morphologically complex, the 'affix', namely the SA, can be regarded as 'inflectional affix writ

large'. All the major syntactic properties of SAs follow naturally from this analysis.

In the next section, I review Chomsky's (2000, 2001, 2004) influential Agree theory and its motivations. In section 2, I present my main proposal, which is an extension of Agree theory. In section 3, I discuss how this proposal can naturally account for (1). Section 4 concludes the chapter.

4.1 The Agree theory

Agree theory (Chomsky 2000 *et seq.*) refines the minimalist theory of movement. The main idea is that syntactic dependencies can be established by a featural operation involving two syntactic objects in a local c-command relation. The motivation for this theory is theory-internal. Basically, it is driven by Ockham's Razor, to eliminate as many unnecessary technical mechanisms as possible. More specifically, it is motivated by the inadequacy of Checking theory and Attract theory, which involves non-minimalist devices such as spec-head configuration and feature movement, respectively. This theory is influential in that in deals with displacement phenomena in a rather minimalist fashion, with no more theoretical constructs from GB-era phrase structure rules. Another major consequence of this theory is the reinterpretation of the nature of syntax-morphology interface that extends beyond φ-feature agreement (Watanabe 2004, Zeijlstra 2004, Penka 2007, Pesetsky and Torrego 2007, Haegeman and Lohndal 2010, etc). It is mainly this consequence of the Agree theory that I will address in this section, although the others are also relevant.

4.1.1 The definition

In the minimalist framework, a lexical item enters the numeration with either interpretable or uninterpretable features (Chomsky 1995: 277). Uninterpretable features enter derivations without values, while interpretable features carry values from the outset (Chomsky 2001: 5). The values of uninterpretable features are determined by Agree. After Agree occurs features must be deleted from the narrow syntax but left available for the phonology.[1] In addition, the syntactic configuration of Agree involves locality requirements (Chomsky 2000: 122):

---

[1] In the literature of morphology and morphosyntax, the terms 'controller (the element which determines the agreement)' and 'target (the element whose form is determined by agreement)' are sometimes used (Corbett 1998, Baker 2008, Bobaljik 2008, etc). Under the present framework, controllers are expressions that bear valued interpretable features, and targets are those that bear unvalued uninterpretable features and get values from the former or the pied-piped XPs that contains X^0 with such features. I will use these terms liberally through this chapter.

(2)   a. Goal G must (at least) be in the *domain* D(P) of probe P.

b. D(P) is the sister of P.

c. Locality reduces to "closest c-command."

Here *probe* is the member of the Agree relationship that enters the derivation later, and is therefore the higher member. *Goal* is the member that enters the derivation earlier, and is structurally lower. Finally, there are two general principles with respect to the 'activator' of the Agree operation:

(3)   a. P is the element that activates Agree. (Chomsky 2001: 5)

b. Uninterpretable features serve to implement operations. (Chomsky 2000: 123)

Based on these two principles, it follows that unvalued uninterpretable features must be carried by the probe. This and certain properties of structural case lead to the proposal of the "Activity Condition": valuation of case features is not an Agree operation and is merely a by-product of φ-feature agreement. This view of Agree runs afoul of certain empirical facts, such as subjects with quirky cases (Nevins 2005), and is also challenged on conceptual grounds by Pesetsky and Torrego (2007). In what follows I will assume there is no Activity Condition[2] and no such principles as (3), and either probe or goal may carry unvalued uninterpretable features in an Agree operation.[3]

4.1.2 Inflectional morphology and syntax-morphology interface

According to Chomsky (2001: 5, 2004: 116), the simplest assumption about an uninterpretable feature F is that it enters the derivation without value. This is so because the value is determined only in the syntactic context by Agree. For example, in the sentence *John laughed* the inflected verb *laughed* is uninflected in the numeration, and is only inflected after Agree applies. This view of inflectional morphology departs from Chomsky's (1995, ch3) analysis of inflectional morphology under Checking theory in that there he assumes *laughed* enters the derivation fully inflected.[4] For convenience, let's call the former the *bare-form* analysis, and the latter *rich-form* analysis. The (bare-form) Agree analysis implies that the motivations for the rich-form analysis must be in error or in need of reinterpretation.

---

[2] However, a modified Activity Condition may exist, such as Baker's (2008: 155) *Case-Dependency of Agreement Parameter*. Our Goal Condition (8a ii) below also achieves similar effects, despite notable differences.

[3] Adger (2003), Baker (2008), among others, also adopt this view of Agree.

[4] In chapter 4, however, Chomsky weakens this assumption (see p.239). See also chapter 5 of Lasnik (1999b) for criticisms of this approach, which he replaces with a hybrid approach that involves affix-hopping as a PF operation.

There are two main arguments for the rich-form analysis. The first is the principle that raising is preferred to lowering, with the assumption that the bare-form analysis would require Infl lowering to V overtly (Chomsky 1995: 139). One problem with this argument is that a bare-form analysis does not need to involve a lowering analysis. In fact, it has been argued that a bare-form analysis may involve a PF operation (Lasnik 1999b), or the 'phonological component' (Chomsky 2004: 116), both of which do not involve syntactic lowering. In addition, even if some kind of syntactic lowering is involved, it is no longer clear whether lowering is banned in grammar. To see this, consider the original argument against lowering in Chomsky (1995: 139). In the Checking theory, lowering of the Infl to V, as well as any movement, creates a trace t. If t is not c-commanded by the head of the chain at LF, the result is an improper chain. Therefore, [$_V$ V-Inf] has to move back to the position of t to form a proper chain. These derivations are less economical than the rich-form analysis, since if verbs like *laughed* are inflected in the lexicon, we only need LF movement of the verb to the Infl position to check the relevant features. This reasoning, although natural in a framework that has syntactic objects such as traces and chains, is no longer even formulatable in the later versions of generative grammar, where traces and chains lose their theoretical statuses as a result of the inclusiveness condition. More recent theories of grammatical architecture in fact seem inconclusive about the status of lowering. One principle that potentially bans lowering is the definition of Move (Chomsky 2000: 135):

(4)   a. A probe P in the label L of α locates the closest matching G in its domain.
       b. A feature G′ of the label containing G selects a phrase β as a candidate for "pied-piping."
       c. β is merged to a category K.

(4a) in fact presupposes that the unvalued uninterpretable feature [uEPP]/[uOCC] that triggers movement is the higher member of the probe-goal pair. This presupposition is unmotivated, however, since other uninterpretable features such as [$u$wh] on wh-phrases and case features on nouns, are the lower members of the probe-goal pair. It is not clear why only uninterpretable features that trigger movement are required to be the higher member. Another principle that potentially bars lowering is cyclicity of derivation, also known as the Extension Condition. According to the recent formulation of this condition (Chomsky 2004: 117), Merge to α must be at the edge of α. Lowering is barred by this condition if "Merge to α" is defined as any new Merge operation that applies after α is formed. Thus in (5), the lowering operation, Merge to δ, is in fact Merge to α, since it applies after α is formed. Since this Merge is not at the edge of α, the extension condition is violated.

(5)
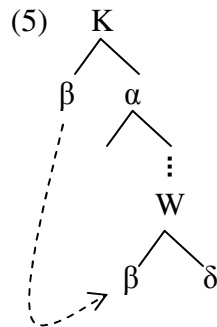
It is not clear, however, whether "Merge to α" is defined as any new Merge operation that applies after α is formed. In the lowering operation such as (5), the operation is more plausibly triggered either by some feature that involves β and δ, but not α. In such a case, it is more precise to regard it as 'Merge to δ.' Thus, it is also not clear that lowering in (5) is barred by the Extension Condition. The result of these considerations is that nothing in the current minimalist framework bars the option of lowering, and that the first motivation doesn't hold water.[5] The second argument for the rich-form analysis is the various considerations that support the Lexical Integrity Hypothesis (LIH), also know as the Atomicity Thesis, with the following proposal (Di Sciullo and Williams 1987: 49):[6]

(6)   Words are 'atomic' at the level of phrasal syntax and phrasal semantics. The words have 'features,' or properties, but these features have no structure, and the relation of these features to the internal composition of the word cannot be relevant in syntax.

The consequence of (6) for inflectional morphology is that since word-formation rules are autonomous, these rules must have applied before syntax rules apply, hence the rich-form analysis. The strongest arguments for LIH, however, do not involve inflectional morphology that involves tense, agreement, and case. These are, however, apparently syntactic, since they are directly involved in Agree operations. Furthermore, under current minimalist view of grammatical architecture, *narrow syntax* (NS), the *phonological component* (Φ), and the *semantic component* (Σ) proceed cyclically in parallel (Chomsky 2001: 5, 2004: 107). A consequence of this for morphology is that the latter can apply in the lexicon as well as be fed by syntactic derivations. From these considerations, arguments for LIH actually have no bearings on the rich-form analysis. We can, therefore, conclude that the bare-form analysis of inflectional

---

[5] Overt lowering analyses have been proposed to account for the syntax of quantification (McCawley 1988), certain VSO languages (Chung 1990, 1998), free inversion in Italian (Rizzi 1982, Burzio 1985), and are associated with right-wrap operations in Categorical Grammar (Chung 1990). It is worth exploring in the future how these analyses can be incorporated within the framework developed here.

[6] Chomsky's (1995: 195) term for LIH is "lexicalist phonology."

morphology à la Chomsky (2001, 2004) is a plausible one.

4.1.3 Pied-pipe and internal Merge

In addition to feature-valuation and deletion, Agree may also trigger other operations. One such operation is generally known as pied-piping or Pied-pipe, which involves determining the size of the constituent that will participate in further operations triggered by Agree (Chomsky 2000: 101). Another related operation is internal Merge (Chomsky 2004: 110). This operation takes the result of pied-piping (which is already present in the derivation at this point) and Merge it again to the edge of the head that bears an uninterpretable EPP/OCC feature, which is also one of the members of original Agree operation. Many questions remain as to the exact nature of pied-piping and the EPP feature, and whether lowering is a legitimate operation (see note 5). In what follows, however, I will address a different kind of problem, namely that Pied-Pipe and internal Merge are insufficient, and one more operation, which is in fact already well-motivated in grammar but has not received due attention, must be added to the arsenal of operations that are triggered by Agree.

4.1.4 Pair-Merge

In the current minimalist framework, adjunction is understood as a process that involves pair-Merge, instead of set-Merge. As we have shown in §2.2.1.2, pair-Merge has several distinctive features: (i) it requires the existence of a 'separate plane'; (ii) this separate plane is a consequence of 'predicate composition', (iii) it requires a SIMPL operation to remove the separate plane and relocate the adjunct to the 'primary plane.' This analysis has conceptual and empirical problems, however. The most salient is that not all adverbial adjuncts can be regarded as complex predicates. Sentence adverbs and focusing adverbs, for example, take scope over the proposition, and do not constitute predicates. Furthermore, the extra machinery of separate planes and the SIMPL operation add operative complexity to the grammar, which is to be avoided if possible. In what follows I will propose a different solution to adjunction that avoids these problems.

4.2 An Agree analysis of sentence adverbs

I propose an analysis for the syntax of focusing adverbs and sentence adverbs, involving two core parts, which I review below.

4.2.1 Agree

First, I propose the basic definition of Agree given in (7):

(7) *Agree*
   a. Match: A feature F (a *probe*) on a head H at syntactic location α searches its c-command
      domain for another F (a *goal*) at location β with which to agree.
   b. Valuation: Replace any unvalued feature with valued feature at α and β.

This is basically Chomsky's (2000, 2001, 2004) version of Agree, departing from (3) in that
unvalued uninterpretable features can be either at the probe or at the goal. There is no need for
both P and G to be *active*.

4.2.2 Agree and focusing adverbs

Next, I adopt the specific analysis of Agree and focusing adverbs given in (8). There are
four basic components: Agree (8a), Pied-pipe (8b), and two instances of Merge (8c,d).

(8) a. (i) X   D/*v*/Aux, etc.   →   X   D/*v*/Aux, etc.
            valued [*i*F]   unvalued [*u*F]   valued [*i*F]   valued [*u*F]
       (ii) Goal Condition: The label bearing the *goal* bears the focus of the *probe*[7]; additionally,
            the main Aux or the main verb may also bear the *goal*.[8]
       (iii)Directionality: The head bearing [*i*F] c-commands the head bearing [*u*F].[9]
     b. A feature [*u*F]' of the label containing [*u*F] selects a phrase P(*u*F) as a candidate for
        pied-piping.[10]
     c. Select a suitable expression M and merge it to the edge of P(*u*F). M realizes the feature
        valuation of [*u*F].[11] The syntactic category of M is A (Adjective/Adverb).
     d. An EPP feature at X triggers internal Merge of [M P(*u*F)] to edge of XP.[12]

---

[7] See Miyagawa (2010) for a similar proposal about *wh*-movement.
[8] Note this implies neither a focused expression nor the main Aux/verb necessarily induces intervention effect so
that a closer potential target always blocks a more remote target. Examples like *John will even leave* and *John likes
only Mary* show that the intervention effect is indeed absent. This may suggest a multiple-Agree analysis or require
some relativiation of the locality principle in the Agree operation (cf. Chomsky 2004). I will leave the details aside.
[9] See also Baker (2008: ch5) for discussion of 'the directionality of agreement' parameter.
[10] Here I follow Chomsky's formulation of pied-piping (4b). I leave the nature of [*u*F]' for future research.
[11] M can therefore be regarded as a *realizer* in the sense of Katzir (2011).
[12] Presumably there are some qualifications of this internal Merge: (i) [M P(*u*F)] undergoes internal Merge if and
only if P(*u*F) is the focus of the probe. (ii) if M is not attached to the focus of the probe, then only the focus
undergoes internal Merge covertly. I will leave these details aside.

e. Direction of merger is determined in the Φ.

Agree occurs between the probe, a valued interpretable feature [F] of a higher functional head, and the goal, an unvalued uninterpretable [uF] of a structurally lower functional or lexical head (8a). In the case of an adverb like *only*, for example, we might specify [F] as the feature [Id] for "identification",[13] and X as the functional head that hosts [Id]. Thus a valued interpretable [Id] of X searches X's c-command domain for an unvalued uninterpretable [uId] hosted by a structurally lower functional or lexical head, and agrees with it, etc.[14] An [uId]' feature of the label containing [uId] (the goal) enables a domain for pied-piping P([uF]), which is either the [F]-bearing head itself or some projection of it (8b). A morphosyntactic expression M is then selected and internally merged to P([uF]) at its edge. The entire [M P([uF])] complex is then raised to the probe's specifier position, either overtly or covertly (8c,d).

Applied to sentence adverbs in particular, the general picture remains the same, with the distinctive aspect being that the probe is located at C^0 and the relevant feature for Agree is [Mood] (9):

(9) *An Agree analysis of sentence adverbs*
   a. (i) C   D/*v*/Aux, etc.   → C   D/*v*/Aux, etc.
         valued [iMood] unvalued [uMood]   valued [iMood] valued [uMood]
      (ii) Goal Condition: The label bearing the goal bears the focus of the probe; additionally, the main Aux or the main verb may also bear the *goal*.
      (iii)Directionality: The head bearing [iMood] c-commands the head bearing [uMood].
   b. A feature [uMood]' of the label containing [uMood] selects a phrase P([uMood]) as a candidate for pied-piping.
   c. Select a suitable expression M and merge it to the edge of P([uMood]). M realizes the feature valuation of [uMood]. The syntactic category of M is A (Adjective/Adverb).
   d. An EPP feature at C triggers internal Merge of [M P([uMood])] to edge of CP.
   e. Direction of merger is determined in the Φ.

---

[13] See Horvath (2007), who proposes an [EI] (exhaustive identification) feature for *only*. I assume here that [Id] is the feature and [EI] is the value.
[14] To understand condition (8a ii) will require understanding why uninterpretable features exist in the first place, which is still an obscure proposition. One possible idea is to associate it with Baker's (2008: 155) *Case-Dependency of Agreement Parameter*:

(i)   F agrees with DP/NP only if F values the case feature of DP/NP or vice versa.

Presumably, focus plays the same function as case, so (i) and (8a ii) can be derived from the same principle. I will leave this issue open.

Note that in this proposal the first two operations (8a,b)/(9a,b) are not new; they are simply components of the operation Move. The third operation (8c)/(9c) is new. The domain of pied-piping P(*u*F) does not undergo Merge *at another location*, but instead undergoes Merge at the same position with a new syntactic object M. This novel Merge operation is distinct from the known external Merge and internal Merge in that *it inserts something new within a syntactic object that has already been generated*. It is also distinct from Chomsky's (2004) version of pair-Merge in that it doesn't resort to the notions of 'separate plane' and 'primary plane', or the SIMPL operation. Its function is also distinct from the known Merge operations: it morphosyntactically *realizes* the result of the feature valuation produced by Agree.[15]
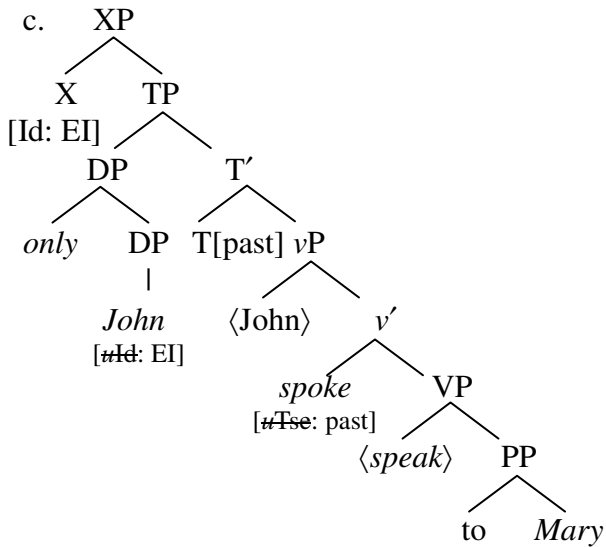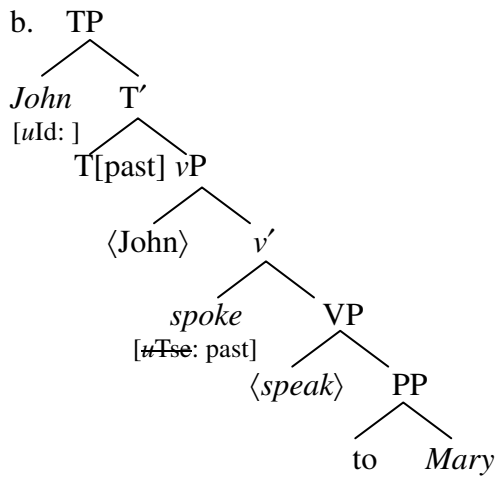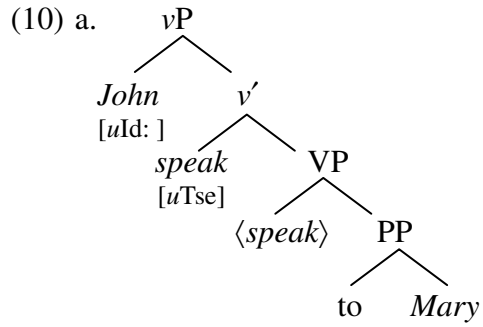
The Merge operation (8c)/(9c) might be intuitively thought of as *delayed-Merge*, in that the operation applies within a syntactic object that has already been generated. The apparent countercyclic nature of the operation may seem problematic, but if we think about the overall architecture of grammar in the Agree theory, the problem is seen to be only apparent. The Agree operation allows unvalued uninterpretable features such as case features on D and tense features on V to remain temporarily unvalued while the syntax builds structures. These features are valued and deleted only after the probe T is merged. This means that feature valuation and deletion is not required to occur immediately; they can be delayed. All things being equal it thus follows that other syntactic operations, such as Merge, should not be required to occur immediately either. If we can delay agreement with the goal until a point where higher structure has been formed, we should be able to delay merger with the label containing the goal until a point where higher structure has been formed, etc.[16]

4.2.3 Some derivations of focusing adverbs

Let us consider some detailed derivations illustrating this analysis, beginning with examples involving simple focusing adverbs like *only*. Under the proposals in (7) and (8), the sentence *Only John spoke to Mary* has the derivational stages shown in (10a-c):

---

[15] This also departs from Chomsky's (2004) view of the function of pair-Merge, which he assumes to be predicate composition.

[16] Contra Chomsky (2004: 117), which states "non-cyclic Merge to a term properly contained in α complicates all three parallel derivations: NS, Φ, Σ." If Agree (and delayed valuation) is an unavoidable operation, it could presumably simplify grammar, even as a best possible solution, instead of complicating it, according to Chomsky's SMT tenet.

(10) a.

```
            vP
           /  \
       John     v′
      [uId: ]  /  \
           speak    VP
          [uTse]   /  \
              ⟨speak⟩   PP
                       /  \
                      to    Mary
```

b.

```
          TP
         /  \
     John     T′
    [uId: ]  /  \
        T[past]   vP
                 /  \
           ⟨John⟩     v′
                     /  \
                 spoke    VP
               [u̶T̶s̶e̶: past] /  \
                      ⟨speak⟩   PP
                               /  \
                              to    Mary
```

c.

```
            XP
           /  \
          X     TP
       [Id: EI] /  \
             DP      T′
            /  \    /  \
        only    DP  T[past]  vP
                |          /  \
              John    ⟨John⟩     v′
          [u̶I̶d̶: EI]           /  \
                          spoke    VP
                       [u̶T̶s̶e̶: past] /  \
                                ⟨speak⟩   PP
                                         /  \
                                        to    Mary
```

(10a) represents the stage when *v*P has been formed. Here, the V+*v* complex bears several uninterpretable, unvalued features, including (among others) [*u*Tense] (abbreviated as [*u*Tse]). The subject *John* bears the [*u*Identification] feature (abbreviated as [*u*Id]) associated with
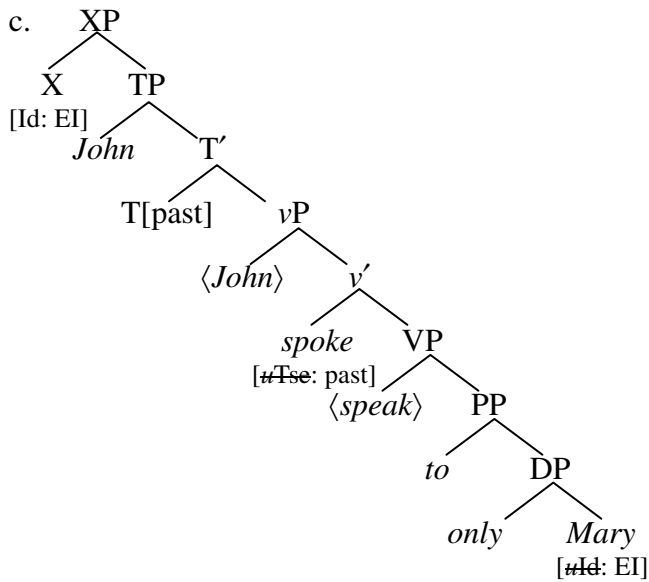
174

focus.[17] (10b) represents the stage TP where has been formed, the [*u*Tse] on the noun being valued by Agree with the corresponding feature at T; the appropriate morphological rule applies, deleting the feature. In (10c), the head X that hosts the valued interpretable [Id: Exhaustive Identification] (abbreviated [Id: EI]) is merged to TP. Agree between [Id] at X and [*u*Id] at D provides value for the latter. An [*u*Id]' feature then selects the DP as the domain of pied-piping. The focusing adverb *only* then undergoes delayed-Merge with the DP, deleting the [*u*Id] feature.

The surface order of *Only John spoke to Mary* does not reveal whether the constituent *only John* has moved to XP spec overtly but vacuously, or covertly as allowed by (8d). To settle this consider next the sentence *John spoke to only Mary, which* has the derivational stages shown in (11):[18]
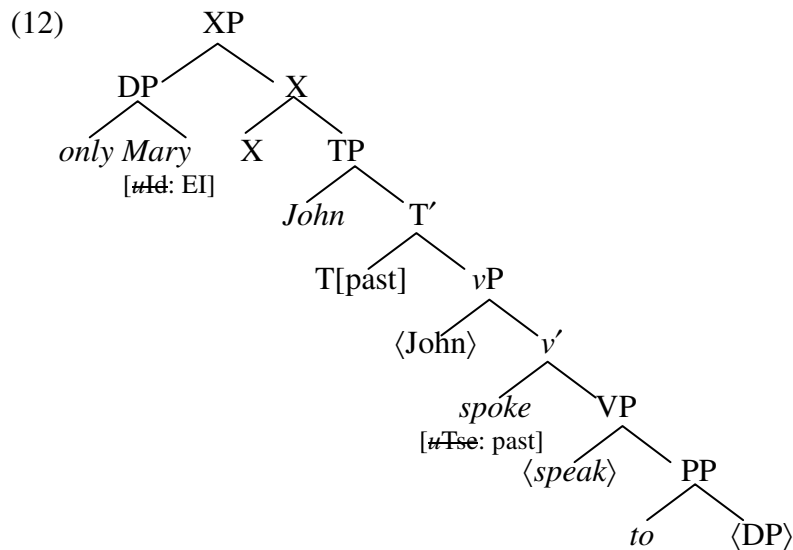
(11) a.

b.

---
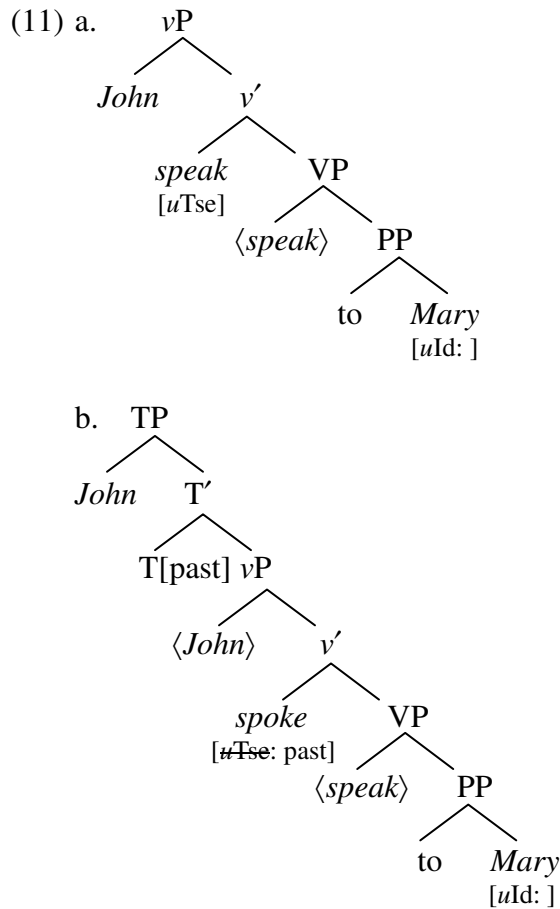
[17] I will abstract away from the internal make-ups of noun phrases. See Chomsky (2007) for a recent discussion.

[18] For expository reasons, some features and structures are omitted, such as the EPP, φ, and case-related features and the details of the V to *v* movement. Head movement is a problematic mechanism in recent versions of generative grammar (cf. Fukui and Takano 1998, Chomsky 2000, 2001, Matushansky 2006 etc). I will keep neutral about its proper analysis.

c.

```
          XP
         /  \
        X    TP
      [Id: EI] / \
           John   T′
                 /  \
            T[past]  vP
                    /  \
              ⟨John⟩   v′
                      /  \
                 spoke    VP
               [uTse: past] / \
                     ⟨speak⟩   PP
                              /  \
                            to    DP
                                 /  \
                             only    Mary
                                   [uId: EI]
```

Again (11a) represents the stage when *v*P is formed, with V+*v* bearing [*u*Tse], and *Mary* bearing [*u*Id]. (11b) represents the TP stage. (11c), XP, the projection of X⁰ which hosts the valued interpretable [Id: EI] feature, is added. Agree between [Id] at X and [*u*Id] on *Mary* provides value [EI] to the [*u*Id] feature on the object. After valuation, an [*u*Id]' feature selects DP as the domain for pied-piping, and the focusing adverb *only* undergoes delayed-Merge with DP, deleting the [*u*Id] feature and realizing the feature valuation. The fact that we pronounce this sentence *John saw only Mary* versus *Only Mary John saw* indicates that movement of the [M P([*u*F])] is covert in English. Hence the final representational stage for this sentence is appropriately (12):

(12)

```
              XP
             /  \
           DP    X
          /  \  / \
   only Mary  X   TP
   [uId: EI]     / \
            John   T′
                  /  \
             T[past]  vP
                     /  \
               ⟨John⟩   v′
                       /  \
                  spoke    VP
                [uTse: past] / \
                      ⟨speak⟩   PP
                               /  \
                             to    ⟨DP⟩
```
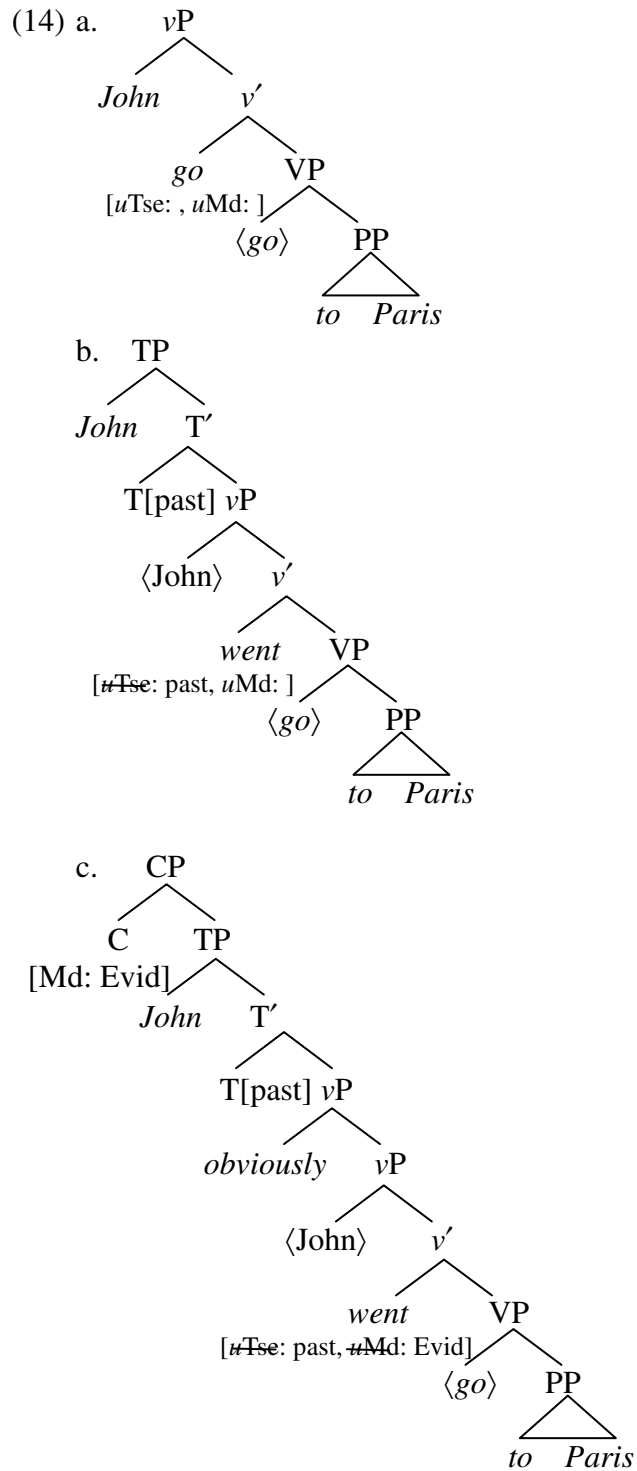
Consider next the sentence *John only saw Mary*, where *Mary* is the focus of *only*. (13a) represents the stage when *v*P is formed, with V+*v* bearing [*u*Tse] and [*u*Id] and Mary bearing [*u*Id]. Here we have a case of multiple-goal configuration, since both a verbal head and a nominal head bear an [*u*Id] feature. This configuration is allowed by (8a ii), and is needed because *v*P is the host for the adverb and the object DP is the focus and undergoes (covert) movement. In (13b), XP is formed, with Agree between [Id] at X and [*u*Id] on *saw* and [*u*Id] on *Mary* providing value for the goals. After valuation, an [*u*Id]' feature on the verb *saw* selects *v*P as the domain for pied-piping, and the focusing adverb *only* undergoes delayed-Merge with *v*P, deleting both of the [*u*Id] features and realizing the feature valuation.[19] Triggered by an [EPP] feature, *Mary* undergoes covert movement to the edge of XP. I assume that Φ determines which of the two potential hosts (*v*P and DP) accommodates the focusing adverb.

(13) a.

b.
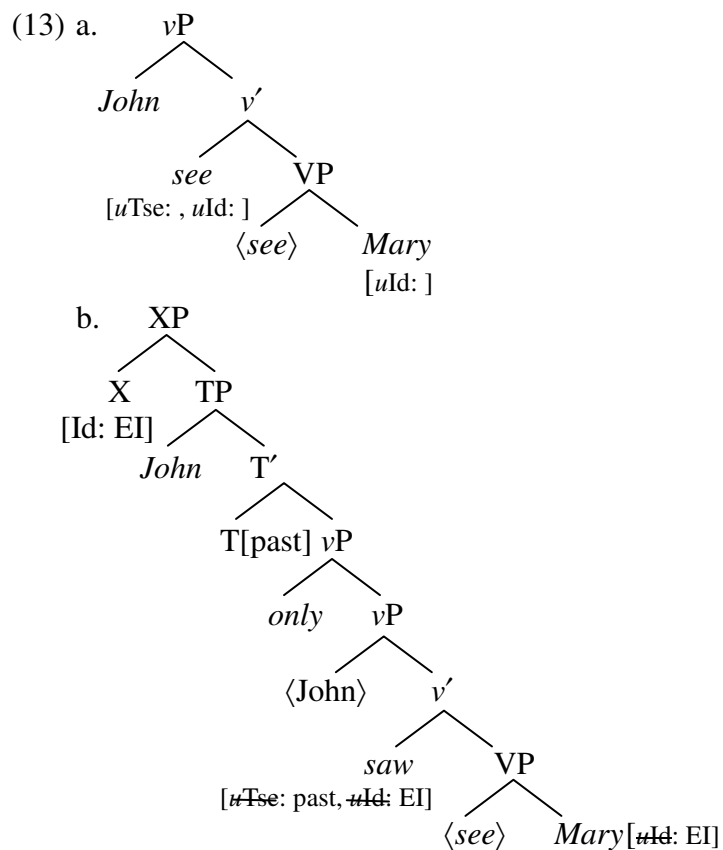
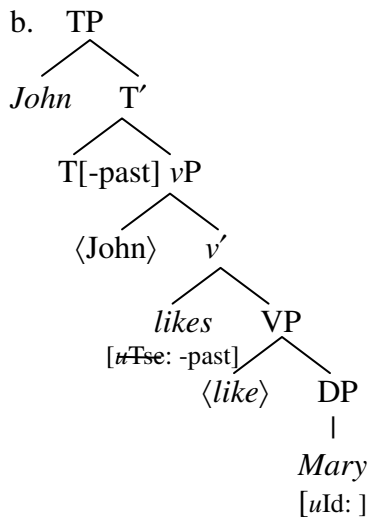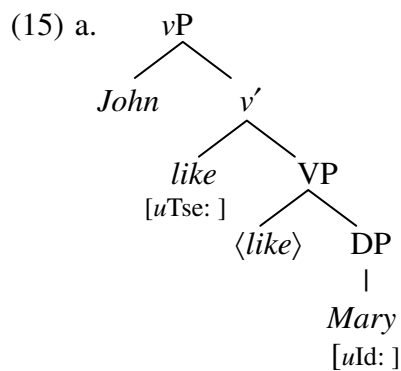[19] It has been suggested to me by Dan Finer and Richard Larson during my thesis defense that only *Mary* is the locus of the [*u*Id] feature, and *v*P is selected for pied-piping, resulting *v*P being the host of the FA *only*. However, it is not clear this single-goal analysis is superior. In fact, *v*P can be very 'large', such as in *John **only** requested that students learn English*F. A single-goal analysis will have to allow very large-scale pied-piping (crossing a clause-boundary), which seems too powerful to me. Furthermore, as I will show below, to allow the verb as the locus of the *goal* can account for a number of syntactic facts by independently motivated economy/locality principles. Pending a better understanding of pied-piping, which deserves a separate research topic, I will stay with the multiple-goal analysis. See also §4.2.5 for more discussion.
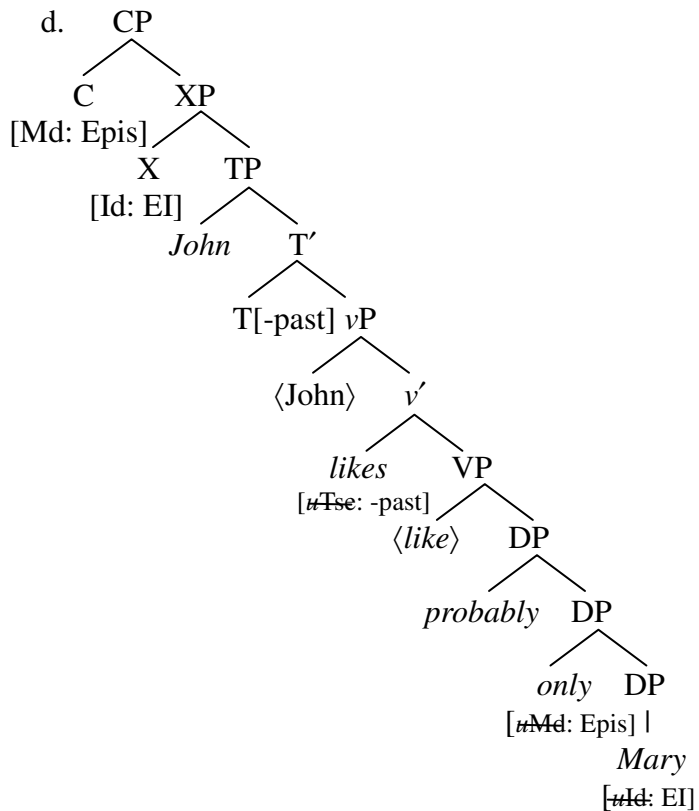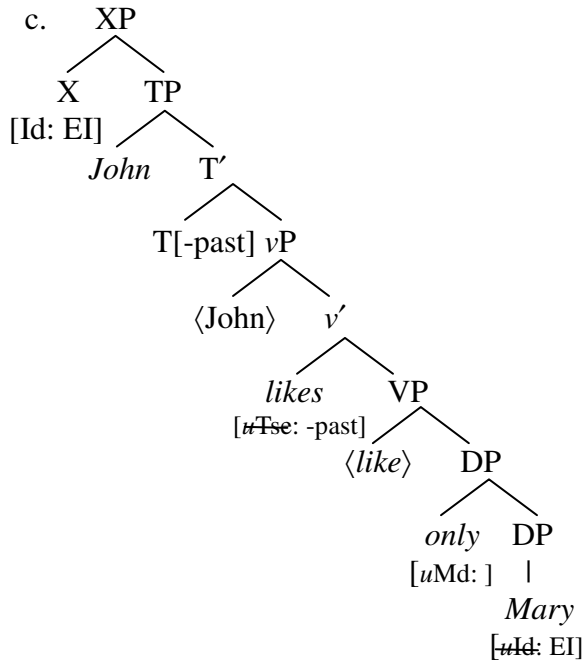
Next consider sentences with sentence adverbs. Since the probe of sentence adverbs is at C^0 but all the other properties are just like typical focusing adverbs, the derivations are quite similar to those we saw above. A sentence like *John obviously went to Paris* has the following syntactic derivations:

(14) a.

```
          vP
         /  \
      John    v′
             /  \
           go    VP
     [uTse: , uMd: ]  \
                ⟨go⟩   PP
                      /  \
                    to   Paris
```

b.

```
        TP
       /  \
    John   T′
          /  \
      T[past]  vP
              /  \
          ⟨John⟩  v′
                 /  \
              went   VP
        [u̶T̶s̶e̶: past, uMd: ]  \
                      ⟨go⟩   PP
                            /  \
                          to   Paris
```

c.

```
           CP
          /  \
         C    TP
    [Md: Evid]  \
           John   T′
                 /  \
             T[past]  vP
                     /  \
              obviously  vP
                        /  \
                    ⟨John⟩  v′
                           /  \
                        went   VP
              [u̶T̶s̶e̶: past, u̶M̶d: Evid]  \
                              ⟨go⟩   PP
                                    /  \
                                  to   Paris
```

178

(14a) represents the stage at which point *v*P is formed. Here the V+*v* complex hosts the unvalued uninterpretable [*u*Tse] and [*u*Md] (short for [*u*Mood]) features. In (14b), T is merged everything proceeds as before. In (14c), a C with valued interpretable [Md: Evid] (*Evid* stands for *Evidential*) is merged with TP. Agree between this feature and the corresponding feature at V+*v* complex gives value to the latter and then [*u*Md] feature selects the whole *v*P as the object for pied-piping, and the sentence adverb *obviously* undergoes delayed-Merge with the resultant *v*P, deleting the [*u*Md] feature and realizing the feature valuation.

Finally, let's consider cases where a sentence adverb is attached to a noun phrase, which already contains a focusing adverb: *John likes probably only Mary*.

(15) a.

```
        vP
       /  \
    John   v′
          /  \
       like   VP
    [uTse: ]  /  \
         ⟨like⟩   DP
                  |
                 Mary
                [uId: ]
```

b.

```
       TP
      /  \
   John   T′
         /  \
    T[-past]  vP
             /  \
        ⟨John⟩   v′
                /  \
            likes   VP
         [uTse: -past]  /  \
                   ⟨like⟩   DP
                            |
                           Mary
                          [uId: ]
```

c.

XP
├─ X [Id: EI]
└─ TP
   ├─ *John*
   └─ T′
      ├─ T[-past]
      └─ *v*P
         ├─ ⟨John⟩
         └─ *v*′
            ├─ *likes* [*u*~~Tse~~: -past]
            └─ VP
               ├─ ⟨*like*⟩
               └─ DP
                  ├─ *only* [*u*Md:]
                  └─ DP
                     └─ *Mary* [~~*u*Id~~: EI]

d.

CP
├─ C [Md: Epis]
└─ XP
   ├─ X [Id: EI]
   └─ TP
      ├─ *John*
      └─ T′
         ├─ T[-past]
         └─ *v*P
            ├─ ⟨John⟩
            └─ *v*′
               ├─ *likes* [*u*~~Tse~~: -past]
               └─ VP
                  ├─ ⟨*like*⟩
                  └─ DP
                     ├─ *probably*
                     └─ DP
                        ├─ *only* [*u*~~Md~~: Epis]
                        └─ DP
                           └─ *Mary* [~~*u*Id~~: EI]

In this sentence, there are two focusing adverbs, and one of them is the focus of another. As shown in (15c,d), it is one of the DP components, *only*, that bears the unvalued uninterpretable feature [*u*Md]. Agree between the [Md] feature at C and [*u*Md] at *only* provides value for the

latter and induces pied-piping and delayed Merge of *probably* to [DP *only Mary*]. This derivation has some important consequences, to which I will come back later.

4.2.4 Inflectional affix writ large and parallelism of NS, Φ, and Σ

Our analyses (7-9) have theoretical consequences for how morphology and syntax interact during the syntactic derivations. The syntactic status of focusing adverbs and sentence adverbs is very similar to inflectional affixes, so the former can be regarded as 'inflectional affixes writ large'.[20] Focusing adverbs and sentence adverbs are like inflectional affixes in that (i) both of them are triggered by uninterpretable features undergoing Agree with interpretable features at different syntactic heads, (ii) both of them enter the derivations late. On the other hand, focusing/sentence adverbs are somewhat 'larger' then typical inflectional affixes in that (i) focusing and sentence adverbs often (but not always) have properties of words instead of bound affixes (they can be polysyllabic, can undergo movements, etc.) and (ii) they generally attach to phrases instead of words. These properties follow naturally from the fact that bound affixes are partially produced by some morphosyntactic operation, we can call it Inflect[21], while focusing and sentence adverbs are produced purely by syntactic operations. Their slightly different derivations are shown as follows:

(16) *Derivations of inflectional affixes*   *Derivations of focusing/sentence adverbs*
       Agree   Agree
       Inflect   Pied-pipe
                                                          Delayed-Merge

(16) shows that morphological and syntactic operations can apply at the same point in the derivations, and which one should apply depends on the feature and the value of the feature. This fits the current view of grammatical architecture according to which NS, Φ, and Σ proceed in parallel.

4.2.5 A note on pied-piping

(8b)/(9b) said very little precisely how pied-piping works. To provide a comprehensive

---

[20] This must be distinguished from the notion of *phrasal affixes*, which refers to bound affixes that attach to phrases instead of words (e.g. Saxon genitive -'s). These affixes are morphologically dependent, unlike adverbs, which are not.

[21] Some plausible formulations can be found in Adger (2003: 170-171).

treatment of pied-piping would be beyond the scope of this work. I will, however, to anticipate some of the later, more detailed discussions, briefly sketch the kinds of pied-piping involved in the syntax of focusing adverbs based on our Agree analysis and general observations. Generally, I assume that the following are the possible cases of pied-piping:

(17) Suppose the goal [uF] is on the head α, P(uF) is the pied-piped phrase, then:
    a. if α is not a T, then P(uF) is the maximal projection headed by α.
    b. if α is an FA attached to a phrase β, then P(uF) is [α β].
    c. if α is the head of a specifier of a nominal constituent YP, and if YP is not itself a specifier
       of another nominal constituent, then P(uF) is YP.
    d. if α is T^0, then P(uF) is T^0.
    e. if the lexical specification of an FA allows TP as a host, then P(uF) is the minimal TP that
       contains α.

(17a-d) cover the typical cases of pied-piping, and are not unfamiliar in the literature.[22] Heads, specifiers, and certain adverbial adjuncts are known as typical 'pied-pipers'. The following examples illustrate these cases, with the underlined expressions are pied-pipers:

(18) a. John likes only [this book].
    b. John even [designed a dress].
    c. Sam wrote only [in his room].
    d. Sam probably [only saw Mary].
    e. Bill accepted only [Mary's invitation].
    f. Bill [can] only play piano.

In (18a), the head of DP, *this*, bears the *goal*, the [uId] feature. The whole DP is selected for pied-piping to merge with *only*. In the same fashion, the whole VP and PP are selected for pied-piping in (18b,c), respectively. In (18d), the FA *only* is the head that bears the [uMd] feature and is adjoined to the VP *saw Mary* in an earlier stage of the derivation. The FA *only* and its host VP are selected for pied-piping to merge with *probably*. In (18e), the [uId] feature is on *Mary's*, which is in the specifier position of a DP. The whole DP is selected for pied-piping. In (18f), the auxiliary verb bears the [uId] feature. Here only the auxiliary verb itself is selected for pied-piping.[23]

---

[22] See Horvath (2006) for a recent overview.
[23] Recall our discussion of this analysis in §2.2.1.1. Williams (1994: 192), based on P&P framework, accounts for a similar situation with regard to English negation particle *not* with the following subcategorization specifications:

(17e) depicts non-typical cases, and is illustrated in (19):

(19) a. [John saw Bill], even/too.

    b. Usually$_1$, [only$_2$ female students are very diligent].

In (19a), the noun *John* bears the *goal*, which selects the whole TP for pied-piping to merge with *even/too*. This is induced by the lexical properties of *even* and *too*. In (19b), an examples that comes from our discussion in §3.1.2.3.1, the probe on *only* also selects the TP for pied-piping. Only this time it is not just the lexical specification of *usually* that induces TP-pied-piping, it is the combination of both the position of the *goal* (at the subject position) and the lexical property of the FA *usually* that selects TP-pied-piping instead of VP-pied-piping.

    Note that I implicitly assume that pied-piping as stated in (17) is in general constrained so that the pied-piped phrase P(uF) cannot be too large (see also note 19). In general, the pied-piped phrase is the focus of the *probe*, or a maximal projection that contains the focus of the *probe*. When an FA is attached to the main Aux or main VP but the latter is not part of the focus of the *probe* (e.g. the focus is the subject or object DP), the [uF] feature on the focus cannot select the main Aux or the main VP for pied-piping. This is because the main Aux or the main verb would be too far away for pied-piping according to (17). Let's illustrate this situation with another example:

(20) ⌜John [could] only have been dating <u>Mary</u>⌝. (He couldn't have been dating other girls.)

In (20), the FA *only* is merged with the main auxiliary verb *could*, but the focus is *Mary*. According to (17), the [uId] feature on *Mary* cannot select *could* as the phrase for pied-piping, since the latter is too far away from the pied-piper. Assuming we are on the right track to constrain the 'size' of a pied-piped phrase, we can account for (20) by (i) allowing the main Aux to be the locus of the [uId] feature even when they are not part of the focus, or (ii) by treating the main Aux to be the locus of the [uId] feature because the main Aux is part of the *secondary focus* (cf. §3.1.2.2.2). In either case, we presumably have something like a multiple-Agree configuration (see note 8 and the analysis for (13) above) instead of large-scale pied-piping.

(i) not: __XP[-tense]

      V[+aux, +tense]__

Although this analysis couched in pre-Minimalistist terms, it is similar in spirit to our analyses of FAs and the generalization (17d).

4.2.6 A note on other types of adverbial adjuncts

Since focusing and sentence adverbs share many morphosyntactic properties with other types of adverbial adjuncts, we expect our analysis above also to some extent apply to other adverbial adjuncts. Disregarding for now adverbs with parenthetical prosodic and syntactic properties, I assume the following principle:

(21) Adverbial adjuncts are derived from Agree.

For instance, a *v*P-level adverb will have the following derivations:

(22) a. X   *v* $\rightarrow$ X   *v*
   valued [*i*F]   unvalued [*u*F]   valued [*i*F]   valued [*u*F]
  b. [*u*F]' selects a phrase P(*u*F) as a candidate for pied-piping.
  c. Select a suitable expression M and merge it to the edge of P(*u*F). M realizes the feature valuation of [*u*F].

X in (22a) is a verbal functional or lexical head, its relationship with *v* is presumably like complex predication ('predicate composition' in Chomsky's (2004) term). Merge X to *v* triggers Agree and delayed-Merge of M to *v* or *v*P, hence the resultant adverbial adjunct properties of M. This process is like word-level compounding in that the operation involve two adjacent heads, so it can be regarded as 'compounding writ large.' I will not go into details here and leave analyses of those adverbs for another occasion.

4.2.7 The cartography of syntactic structures

Our analyses also have consequences for theories that deal with the nature of the fine-grained syntactic structures. One immediate consequence is it reaffirms that adverbial adjuncts is regulated by *syntactic* features and operations just like non-adjuncts do, instead of being regulated by semantic considerations alone. This result means it belongs to the 'syntax camp' rather than the 'semantics camp' of theories of adverbial syntax (see §2.2.2). It follows sentence adverbs should play an essential role in the cartography of syntactic structures.[24]

---

[24] The cartography project is an ongoing major line of research in generative grammar (Cinque 2002, 2006, Rizzi 2004, Belletti 2004, Beninca and Munaro, to appear), which, as described by Cinque and Rizzi (2008), "the attempt to draw maps as precise and detailed as possible of syntactic configurations", and by nature is not even a framework or hypothesis but a research topic. It is not clear what the semantics camp has to say about this project.

Another major consequence of our analyses for the cartography project is that the cartography of syntactic structures involving sentence adverbs is the mapping of two domains: (i) the CP domain, which bears the [iMd] feature, (ii) the domain that hosts the [uMd] feature. This dual-domain mapping view departs from most of the earlier works on the cartography project, such as the AdvP-in-Spec and the 'isomorphic' approaches mentioned in previous chapters. Note also our analyses so far have said nothing about what the CP domain looks like in a sentence with multiple sentence adverbs (i.e. Are multiple [iMd] features encoded in a single C^0? Or does each [iMd] feature occupy a separate C^0?), and nothing about how the hierarchy of clausal functional projections are derived in syntax (i.e. Is there a *fixed* universal hierarchy in the sense of Cinque (1999)? Or does the hierarchy depend solely on the semantic properties SEM(H) of the functional heads in the sense of Chomsky (2004)?[25]). I will assume sentences with multiple SAs have multiple C heads which enter the derivations via regular set-Merge, and leave open how the hierarchy of these C heads is derived in syntax.

In sum, in this section we presented our main proposal for the syntactic analysis of focus and sentence adverbs, and suggested a general analysis for adverbial adjuncts in general, all based on the Agree theory, which is independently motivated by minimalist theoretical considerations and robust syntactic facts.

4.3 Analyses of morphosyntactic properties of sentence adverbs

As mentioned at the beginning of this chapter, the following properties of sentence adverbs need to be explained:

(23) a. The syntax-semantics mismatch problem.
b. The theoretical status of adverbial adjuncts.
c. The C^0 properties of sentence adverbs.
d. The syntax of focusing adverbs.
e. Sentence adverbs are a heterogeneous group.
f. Cross-linguistic variation.

In this section, I show that these properties provide strong empirical supports for our analysis due to the fact they can be naturally accounted for by the latter.

---

[25] See also note 19 in chapter 2.

4.3.1 The syntax-semantics mismatch problem

As has been discussed extensively in chapter 1, a major property of sentence adverbs is that they enjoy a wide semantic scope but overtly occur in syntactically lower positions. This property cannot be naturally accounted for by the earlier versions of generative grammar, such as the one that treats spec-head feature checking as the only mechanism for licensing A′-dependencies. Such a theory either has to ignore syntax-semantics mismatch or resort to ill-motivated and highly controversial covert movement or PF lowering analyses. Our analysis, on the other hand, not only acknowledges the syntax-semantics mismatch but provides a theory that is independently motivated by ubiquitous phenomenon inflectional morphology in human language. The syntax-semantics mismatch is due to the following derivations: (i) the existence of the same feature [F] at two different syntactic heads, one is valued and interpretable, the other is unvalued and uninterpretable; (ii) after Match, established by local or long-distance c-command, some overt morphosyntactic or syntactic material is realized at the syntactic constituent where [uF] is located. Crucially, it is the possibility of bare-form analysis and long-distance Agree (both of which are motivated by minimalist considerations) that allows the option of treating syntax-semantics mismatch phenomena in a natural and uniform fashion, which is not possible in previous theories.

4.3.2 The theoretical status of adverbial adjuncts

As mentioned in chapter 2, the theoretical status of many properties of adverbial adjuncts remains unsettled in recent versions of generative grammar. Let's now review these properties:

(24) *Properties of adverbs and adjuncts*

Properties of adverbs	**Properties of adjuncts**
a. Co-occur with APs, VPs, AdvPs, PPs, IPs, CPs, DPs.	g. Do not change the category or bar-level of the constituent they are joined to.
b. Often derived from adjectives via an affix (*-ly* in English, *-a* or *-os* in Greek, *-mente* in Spanish, *-weise* in German).	h. Optional.
	i. Recursive.
	j. Can be left or right adjoined to the target in certain cases.
c. Cannot be stand-alone predicates (license ellipsis, VP-preposing, etc.).	k. Occur more distant from the head than complements.
d. Can be coordinated with other	l. Can attach at different categorial levels.

adverbial expressions.	m. Free word order in certain cases.
e. Generally do not select and aren't selected.	n. Apparent counter-cyclicity.
	o. Do not block agreement.
f. Inflection marking is mostly absent.	p. Display the Condition on Extraction Domains (CED) effect.
	q. Display the weak island effects in some cases.

Property (24a) is derived from the fact that adverbs in general are heads that do not project, presumably due to the fact that their function is realization of feature valuation (see the discussion of (24e) below), and that their merger with other constituents is generally triggered by Agree instead of c-selection.[26] Consider two simple examples that involve *only*:

(25) a. John only likes Mary.

    b. John likes only Mary.

The simplest assumption is that *only* attaches to *v*P in (25a), to DP in (25b). In our Agree analysis, the different attachment possibilities for *only* come from the fact that either D (where focus is located) or *v* can bear an [*u*Id] feature. In an AdvP-in-Spec analysis, the existence of (25b) is unexpected, and has to resort to additional devices such as "minor" functional heads (Bayer 1996, 1999), which have different properties to normal functional heads. Furthermore, this approach doesn't explain why minor functional heads have the properties they do (not projecting, etc.), while our approach can derive this naturally by the fact that they are derived by delayed Merge. In a pair-Merge analysis (Chomsky 2004), (25b) is also unexpected, since the function of adjunction is predicate composition and adjunction should be limited to only predicates.

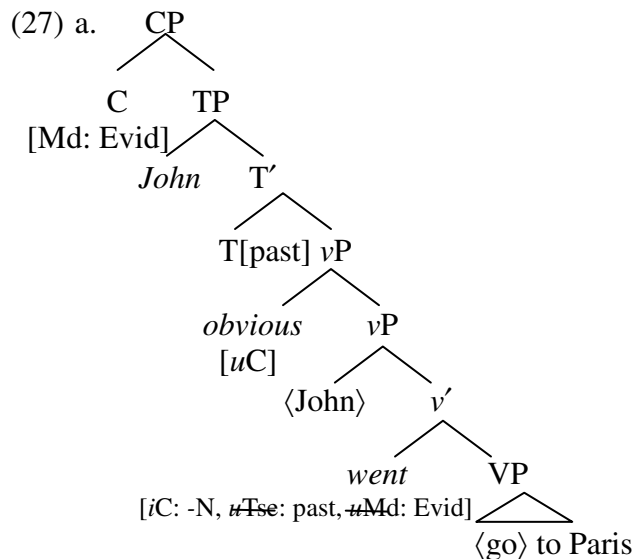(24b) Often derived from adjectives via an affix.

Property (24b), listed above for easy reference, refers to a complex set of facts about adverbial affixes, and it is certainly not a reliable general diagnostic for sentence adverbs cross-linguistically. Nevertheless, it is still important to explain why words like *possible* and *possibly* are in complementary distributions, because the relevant examples are quite robust in English and some other languages. A conceptually attractive analysis would be that adjectives and

---

[26] Sentence adverbs can, in some rare cases, merge with other materials to form a larger 'AdvP' (e.g. *luckily for him*), then the whole constituent merge with its host. The first type of merger is not the same as the delayed-Merge and should receive a separate treatment. See below for categorical properties of adverbs.

adverbs are not separate categories, and their different morphological forms are determined by their syntactic distributions (see Emonds 1985, Radford 1988, Alexiadou 1997, etc. for similar views). I argue this is indeed the case, and argue that that *-ly* is an inflectional suffix, indicating agreement with a [-N] head.[27] More specifically, I propose the following morphosyntactic rule for English:
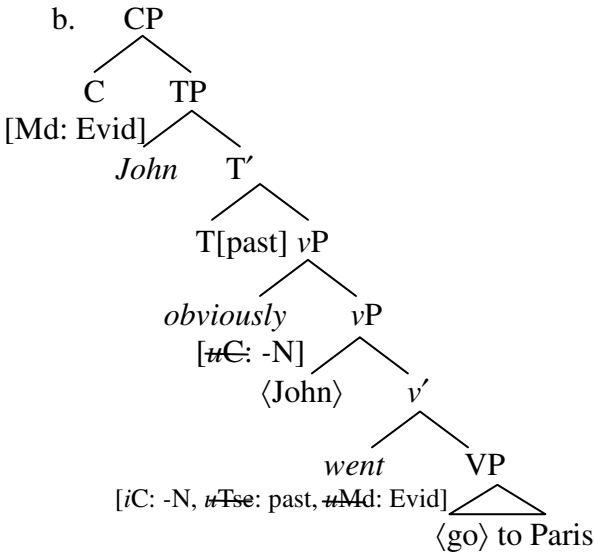
(26) Pronounce A[~~*u*C~~: -N] as *ly*.[28]

This rule states that an expression of syntactic category A (adjective and adverb) is suffixed by *-ly* if its uninterpretable categorical feature C is assigned value [-N], [-N] covering all syntactic categories that are not nouns (e.g. *v*, C, D, P, etc). According to this analysis, a sentence like *John obviously went to Paris* has the following derivations in addition to the ones shown in (14):

(27) a.

```
        CP
       /  \
      C    TP
 [Md: Evid] /\
        John  T′
             /  \
        T[past]  vP
              /    \
         obvious    vP
          [uC]     /  \
              ⟨John⟩   v′
                      /  \
                   went   VP
  [iC: -N, u̶T̶s̶e: past, u̶Md: Evid] /\
                          ⟨go⟩ to Paris
```

b.

```
           CP
          /  \
        C     TP
   [Md: Evid] / \
          John    T′
                 /  \
            T[past]   vP
                     /  \
              obviously   vP
                [uC: -N]  /  \
                    ⟨John⟩    v′
                             /  \
                          went    VP
         [iC: -N, uTse: past, uMd: Evid] / \
                                    ⟨go⟩ to Paris
```

(27a) shows that when the sentence adverb enters the derivation, it has the bare form *obvious*, containing an uninterpretable unvalued [*u*C] feature. It is only after it undergoes Agree with *v* does it get valued and assigned the morphological form *obviously*. When the form is attached to nouns (i.e. *the most obvious example*, *the obvious question*), it gets the value [+N] and therefore do not get the suffix *-ly*. For AdvP-in-Spec and AdjP-in-Spec analysis, we will need to distinguish verbal functional heads and nominal functional heads, etc., in order to derive the correct result. In a pair-Merge analysis, it is not clear why feature valuation of an adverb is possible at all, since when it pair-Merge with the *v*P (or other XPs) the adverb is on a separate plane. It is only after TRANSFER that adverbs can undergo any further syntactic operation. However, if Agree applies within *v*P after TRANSFER, the derivation is counter-cyclic and should be impossible. Thus the presence of the *-ly* suffix again provides support for a delayed-Merge analysis.[29]

(24c) Cannot be stand-alone predicates.

(24c) can be derived from the fact that adverbs do not belong to the syntactic categories that can license ellipsis or VP-preposing. A major descriptive property of VP-ellpisis/VP preposing licensing is that the licensing head is an auxiliary verb. What distinguishes auxiliary verbs from other expressions is their c-selection property: they select *v*P complements. On the other hand, sentence adverbs do not have c-selection properties (although they belong to the syntactic category A). Thus they cannot license ellipse or VP-preposing.

---

[29] It seems plausible that the fact certain adverbial affixes are associated with different scopes of the adverbs (Chinese *-shi*, German *-weise*, Korean *-to*, etc.) can also receive an Agree analysis in which adverbs agree with the controller of their hosts (C or *v*). I will leave those cases for future research.

(24d) Can be coordinated with other adverbial expressions.

(24d) can also be derived from the fact that adverbs belong to the syntactic category A.

(24e) Generally do not select and aren't selected.

(24e) refers to the fact that sentence adverbs and adverbs in general do not select complements or are selected, unlike other syntactic categories. The first property can now be understood as a consequence of the syntactic function of adverbs in general. Unlike other syntactic categories, adverbs enter syntactic derivations not as arguments or predicates, but as morphosyntactic reflexes of feature valuation, just like inflectional affixes. Presumably, some economy consideration requires them to be as morphosyntactically simple as possible. This generally prevents adverbs from any kind of projection. I will not go into details here, but simply point out this property of adverbs can have a principled account under the 'affix writ large' analysis, which is not available in other approaches.[30]

(24f) Inflection marking is mostly absent.

(24f) is like (24c) that can be derived from the fact adverbs belong to the category A, not V. It can only be inflectionally marked when the affix in question can attach to the A category, such as *-ly* and comparative *-er* in English.

(24g) Do not change the category or bar-level of the constituent they are adjoined to.

(24g) is derived from the fact adverbial adjuncts are formed as a reflex of Agree (16). Merge of non-adjuncts, on the other hand, are triggered by c-s related operations. Since adjuncts are reflexes of a syntactic operation, they do not directly take part in feature-checking/elimination process. Non-adjuncts, however, are direct participants of c-s related operations, and directly take part in feature-checking/elimination process. This difference makes adjuncts in effect invisible to any operations (e.g. VP-preposing, co-ordination) that are sensitive to c-s related features. This is a nice theoretical consequence, in that the major property of adjunction is generalized to the nature of Agree operation. Alternative theories either ignore this property of adjunction totally (e.g. AdvP-in-Spec theories), or have to postulate an *ad hoc* distinction between "separate plane" and "primary plane".

---

[30] For Travis (1988), for example, the property that adverbs do not project is a principle itself, not a fact that is derived from some general principles.

(24h) Optional

(24i) Recursive

(24n) Apparent counter-cyclicity

Properties (24h), (24i), and (24n) can also be derived from (16). Like typical inflectional affixes resulting from Agree between their stems and their controllers, if the controller is present, the adverb is present. If the controller is absent, then the adverb is absent. Thus adverbs are optional. When there is more then one controller of the target, we have recursion. They are apparently counter-cyclic because the heads of their hosts can only be valued after theirs controllers are merged, which we have shown to be a consequence of the nature of Agree operations in general.

(24j) Can be left or right adjoined to the target in certain cases.

Property (24j) in fact is quite limited cross-linguistically when it comes to focusing adverbs. For example, the best studied English focusing adverb *only* is in most cases left-adjoined to its host, but it can be right-adjunct to an auxiliary verb (e.g. *He can only play piano*), and in certain cases it can be right-adjoined in cases such as *Passengers only are permitted on the platform*. In addition, certain sentence adverbs in English seem to be able to be right-adjoined to another adverb (e.g. *Pollution will always probably exist*). In general, however, right-adjunction is more restricted in English and in many well-studied languages. I assume that this restriction is mostly determined in the phonology component (8e).[31]

(24k) Occur more distant from the head than complements.

(24l) Can attach at different categorial levels.

Properties (24k) and (24l) are now derived from the availability of pied-piping operation that is independently motivated in grammar. Just as internal-Merge can either involve head-sized or phrase-sized constituents, delayed-Merge can either involve head-sized or phrase-sized constituents to merge with M.[32] According to this view, the fact that adverbs can occur to the right of verbal or adjectival heads (as shown in (28) and (29), taken from chapter 2) is not due to

---

[31] See §4.3.6 for further discussion.

[32] C-s related operations, such as external-Merge, are generally believed to allow only the phrases already constructed to undergo further merger, since the alternatives would induce counter-cyclicity. However, if the ban against strict counter-cyclicity does not hold, as we have suggested in §4.1.2, we should allow a subject NP to occur between a verb and the object NP, where the subject NP is right-adjoined to the verb. This possibility seems to be compatible with the empirical facts of certain VSO languages such as Berber and Chamorro (See Choe 1987, Chung 1990, 1998, who adopt 'lowering' analyses).

head movement, but due to the fact pied-piping doesn't apply and that adverbs (and adjuncts in general) in these cases are right-merged to their host.

(28) a. He isn't proud enough of his country.

    b. The weather may turn out rather frosty.

(29) a. Jean embrasse souvent Marie   (French)

        John embraces often   Mary

        'John often embraces Mary.'

    b. Pierre a   vu   à peine Marie

        Pierre has seen hardly   Mary

        'Pierre has hardly seen Mary.'

    c. Souvent faire mal ses   devoirs, . . .

        Often   make badly Poss homework

        'To frequently do one's homework badly'

Previous right-adjunction analyses (Williams and di Sciullo (1987), Radford (1988), Sportiche (1988), Iatridou (1990), Williams (1994, 2000)) have faced a serious theoretical challenge in that the only motivations for right-adjunction in these cases linguistics could think of is to form complex-predicates and incorporated verbs, which shouldn't apply to sentence adverbs for semantic reasons. Under the present analysis, this theoretical challenge dissolves since Agree and delayed-Merge makes right-adjunction theoretically sound.[33]

(24m) Free word order in certain cases.

Property (24m) involves facts associated with focus-sensitivity and information structure considerations, which I will return to in §4.3.4 below.

(24o) Do not block agreement.

(24o) can be naturally accounted for if we assume the following principle:

---

[33] An additional consequence is that certain cases of V-AdvP-XP sequence in English can be explained in the same way without requiring verb movement, illustrated below (see Pesetsky 1989, Ouhalla 1990, Costa 1996 for motivations for verb movement analysis):

(i) a. Bill knocked recently on it.
    b. Sue looked carefully at him.
    c. Harry relies frequently on it.

(30) An adjunct is headed by an A head.[34]

We have already seen there is evidence that adjectives and adverbs belong to the same syntactic category A. When they occur alone, they do not block (tense and φ-feature) agreement marking on verbs since although the former configurationally intervene between the probe and the goal, they don't have the relevant feature specifications.[35] However, it is also true that adjuncts of other categories, including verbal adjuncts, do not block verbal agreement either:

(31) a. He left the room laughing/crying/singing.
     b. *He leave the room laughing/crying/singing.

Why don't the intervening verbs block Agree in (31a)? A simple analysis to capture the fact that both adverbial adjuncts and VP adjuncts do not block verbal agreement is an analysis based on principle (30). That is, (31a) has the following structure:

(32) He left the room [$_{AP}$[$_{vP}$ laughing]].

The verb *laughing* doesn't block verbal agreement because it's embedded in an AP and does not c-command *left*. The suffix *-ing* would be an inflectional 'adverbialization' suffix, marking agreement between A and *v*. In English, the form is like present participle suffix and nominalization marker, so we don't have direct evidence *-ing* reflects the existence of A. In many other languages, however, there is strong evidence the analysis in (32) is on the right track. This is because the VP adjuncts in sentences like (31a) is marked by specific affixes only associated with adjunction. In German, the form "present participle" (V-*end*) is used for this purpose, which does not have other functions of English *-ing*. In Slavic languages, the form "adverbial participle" is used only for this purpose. In Mongolic and Turkic languages, the form "converb" is only used for this purpose.[36] These morphological facts show that languages can have unique morphological markings to show a VP is used as an adjunct. Following the general assumption that this morphological marking correspond to the syntactic category A, we have the principle (30). It then follows naturally adjuncts do not block verbal agreement, since they are all headed

---

[34] See Rubin (1994, 1996, 2003) for a similar proposal. His argues that the functional head Mod forms an extended projection around all base adjuncts, based on arguments distinct from the ones presented here.
[35] Note that an A (as an adjective) does have φ-feature agreement when it is attached to an NP in certain languages. Their presence never blocks φ-feature agreement on the verb. This is presumably a concord/multiple-Agree phenomenon.
[36] Andrei Antonenko, Katharina Schuhmann, and Aydogan Yanilmaz informed me of the Slavic, German, and Turkic facts, respectively.

by A.

I will postpone discussions of (24p,q) to §4.3.4.

All in all, in light of these analyses of adverbial adjuncts, we can conclude that their theoretical status can be delineated in a way that remains close to the strong minimalist thesis (Chomsky 2004), in that their properties are derived from very general and ubiquitous operations in the grammar of human language: Agree, Pied-piping, and Merge. Principles (26) and (30) are compatible with our analyses and are motivated by independent empirical and theoretical considerations, so they can be naturally incorporated into our analyses. Previous analyses are not able to derive all of these properties in a theoretically coherent way.

4.3.3 The C^0 properties of sentence adverbs

It has been established in chapter 2 that sentence adverbs manifest properties of C^0 elements syntactically, repeated as follows:

(33) a. Ability to scope over the subject of the sentence.
    b. Restricted when under the scope of a clausemate C^0 element.
    c. Selection restrictions with V^0.
    d. Restricted in embedded clauses in other contexts.
    e. Clause-linking function.
    f. Denotation focus and quantification are usually not possible.
    g. Long-distance movement is not possible.

These properties have never been properly formally addressed. Now we can have a natural account without complex *ad hoc* mechanisms such as various head movements and subject movements. The valued interpretable feature [*i*Md] is located at C^0, whereas the unvalued uninterpretable feature [*u*Md] is located at a lower functional or lexical head. The syntax adverbs themselves are licensed by Agree and relevant operations.

Facts associated with in (33a) (see §2.3.2.1) are derived by the presence of a covert C with [*i*Md] feature. Since C^0 c-commands spec-of-TP, it follows it scopes over the subject, and more precisely, the subject NP can be the *focus* of the sentence adverb. When it seems the subject takes wide scope, the sentence adverb in fact still takes wide scope but without taking the subject NP as its focus. When it comes to the focus-sensitivity, this analysis has the further advantage that the overt syntactic position of the sentence adverb can be regulated by a principled account, which we will return to below.

(33b,c,d) can now all in principle be accounted for by c-s related conditions between

various C^0 elements, and between C^0 elements of the embedded clause and the matrix verb. Since sentence with sentence adverbs have a C with [iMd] feature, it is natural this feature is compatible with some features, but not others, on neighboring C^0 elements. This derives (33b). Since only specific clause-types are compatible with verbs taking clausal complements, and that [iMd] feature interact with these clause-typing features due to their proximity, we expect (33c). (33d) describes facts that are more subtle because many embedded clause are not selected by a verb and do not have overt C^0 elements, yet it is plausible that they all have clause-typing C^0 elements that interact with the C that contains [iMd] feature, so the restrictions can be derived in the same way.

(33e) can now also be easily derived from our analysis. Clause-linking expressions involve at least a coordinator or a subordinator, as well as two clausal constituents. The clausal constituents involved have to be of specific moods, encoded by specific C^0 elements bearing [iMd] features, to satisfy the c-s properties of the coordinator or subordinator. This C can either be overtly expressed as a C^0 head, or expressed as a sentence adverb that realizes Agree.

(33f) now naturally follows from the syntactic hierarchy of various focus-sensitive elements, and the generalization/tautology that the focus of an FA must be within its scope. More specifically, sentence adverbs are the product of Agree of C^0 elements that are hierarchically higher than functional heads realized by focusing adverbs such as *only* and *often*, so the former cannot be the foci of the latter. The fact SAs are higher could be due to the existence of a somewhat fixed, somewhat liberal hierarchy of functional categories, or due to semantic constraints, but in any event this seldom-mentioned property of sentence adverbs can now be accounted for in principled syntactic terms.

(33g) is in fact a subcase of (33c) (see §2.3.2.7), and is therefore accounted for in the same way. The first-Merge of a sentence adverb is triggered by the [iMd] feature of the C head of the embedded clause. If the [iMd] feature is incompatible with the verb of the matrix clause, c-selection will not be satisfied and the derivations will crash.

In sum, the various C^0 properties of sentence adverbs now have straightforward accounts under our Agree analysis of sentence adverbs. It is the presence of a covert C bearing [iMd] feature that manifests these properties. Again, there is no way to derive these properties in previous analyses other than proposing *ad hoc* mechanisms such as vacuous movements of non-adverbs, reconstructions, or simply disregarding the problem of syntax-semantics mismatch altogether.

4.3.4 The syntax of focusing adverbs

The property of sentence adverbs extensively discussed in chapter 3 is focus-sensitivity. I

will show all of these properties (Generalizations A-F) can be derived in a principled way from (8) and (9).

4.3.4.1 Generalization A – Three components

(34) The syntax of a focus-sensitive expression involves its dependency relations with its focus, its host, and its scope.

This generalization can be derived straightforwardly from our Agree analysis. The scope position is the position of the controller head with valued [iF] feature, [F] can be a variety of features, including [Md]. The host is the target, which is either the head with [uF] feature, or the phrase P(uF) that is selected for pied-piping. The constituent with focus is one of the possible candidates for the target of agreement, but other targets are also possible.[37] The FA itself is the reflex of Agree, formed by delayed-Merge with the target head or target phrase.

4.3.4.2 Generalization B1 – Free attachment, except to TP

(35) The first-merge host of an FA can be all kinds of syntactic categories an adverb can attach to, except a TP.

This generalization should now be regarded as a consequence of the general property of adverbs (24a) discussed above, and can now be derived in the same way. Let's now examine possible adjunctions sites in more detail:

(36) a. John likes [possibly [$_{DP}$ only Mary]].
　　b. You eat [either [$_{DP}$ all of the ice cream]], or I punish you.
　　c. John [definitely [$_{vP}$ likes Mary]].
　　d. <u>Someone</u> [sure [$_{vP}$ <u>played a prank on someone</u>]].
　　e. John [[$_T$ can] obviously] sing this song.
　　f. You [[$_T$ ought] really] to have stayed.
　　g. [Probably [$_{TP}$ John likes Mary]].
　　h. *[Sure as hell [$_{TP}$ John likes Mary]].

---

[37] The fact that the target of agreement is not completely fixed in a given sentence is not new. Korean plural suffix -*tul* (Kim 1994, etc.) is one such example that comes to mind. Clitic climbing in Romance languages can probably also be analyzed in this way.

In (36a,b), the SAs are attached to the object DP, due to the fact some constituent within DP is the focus of the SAs and can bear [*u*Md] and the consequent Agree, pied-piping and delayed-Merge operations. In (36c,d), the SAs are attached to *v*P, due to the condition (8a ii): the main Aux or the main verb is the default locus of the [*u*Md] feature, and the consequent three operations. In (36e,f), it is the main Aux that bears the [*u*Md] feature. Here, however, no pied-piping applies (cf. above discussions of (24l)). (36g,h) show that while some sentence adverbs can apparently attach to TP, some cannot. This is due to that fact that the TP-initial SA is usually due to the topicalization of the SA itself and that topicalization is constrained by lexical factors (cf. §3.2.2.2).

The various attachment sites thus follow from our analyses (7-9). However, it must be noted that the non-attachment to TP is associated with our proposal of how pied-piping works in (17d) (only T^0 is pied-piped when it is the target), which still requires further investigation.

4.3.4.3 Generalization B2 – Maximal projection host

(37) An FA attaches only to maximal projections.

As we have seen in Chinese and German examples in §3.1.2.2.1, focusing adverbs generally have to attach to maximal projections. This can now be derived from our pied-piping analysis in (17). Although the *goal* is located at a given head, it is a certain given maximal projection that is selected for pied-piping for delay-Merge (and later internal-Merge).

(38) a. …weil man den Wagen **nur** in die Garage <u>fahren</u> darf.
    b. *..weil   man den Wagen   in die Garage **nur** <u>fahren</u> darf.
      Because one   the car   in the garage   only drive   may
      'Because you may **only** <u>drive</u> the car into the garage.'
(39) a. ta **zhi**   wei zhangsan xie-le   <u>yi-ben</u> shu
    b. ??ta   wei zhangsan **zhi**   xie-le   <u>yi-ben</u> shu
       he for Z.   only write-Pfv one-Cl book
     'He **only** wrote <u>one book</u> for Zhangsan.'

In (38) and (39), it is the main verb (raised to *v*) that bears the [*u*Id] feature, courtesy of its being the focus in (38) and its simply being the main verb in (39). The phrase P(*u*Id) they select for pied-piping include the oblique argument or *v*P adjunct since the latter is part of the *v*P projection.

4.3.4.4 Generalization C1 – C-command Condition

(40) Some FAs that are attached to clausal projections (VP, TP, etc.) have to c-command their foci at overt syntax. Generally all FAs attached to non-clausal projections have to c-command their foci at overt syntax.

This generalization follows from the Goal Condition (8a ii), the derivations that lead to delayed-Merge of an FA and a general economy/locality principle that I call Shortest Agree. When the goal is in the head of a non-clausal constituent, it has to bear the focus of the probe. Since delayed-Merge attach FA to the pied-piped phrase P(uF) that contains the goal, FA has to c-command the focus. When the goal is in the head of a clausal constituent, it doesn't need to bear the focus of the probe, so the C-command Condition need not be in effect. However, when the condition does apply, such as in the case of *only*, we need to add the following lexical specification to those FAs:

(41) Intervention parameter: Attach FA to a $F_{(+N)}$ constituent if the latter is the closest potential target to the probe [F].

Generally, Generalization C1 follows from (7-9), while lexical variations such as (*_John will only play piano._ vs. _John will even play piano._) follow from parameter (41). The syntactic dependency between the FA and its focus thus follows not directly from Agree between FA and its focus, but from Agree between a covert operator bearing [F] and a [uF] in a head that bearing the focus, and the resultant delayed-Merge of FA to a pied-piped constituent of the head.

The fact that certain sentence adverbs only occur in a post-subject position unless the subject NP is the focus of these adverbs can now be derived (see §3.2.2.2 and §3.2.2.4) from (7-9), especially the Goal Condition. Consider the following two sentences of Chinese again:
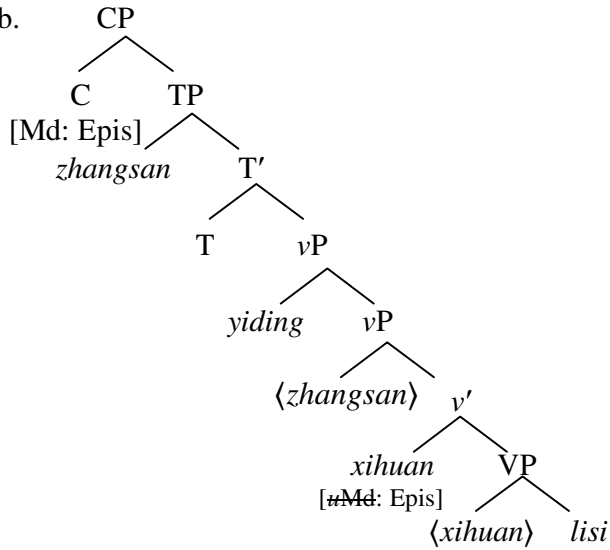
(42) a. (*yiding) zhangsan (yiding) xihuan lisi
           surely   Z.   surely like   L.
         'Zhangsan surely likes Lisi.'
     b. (yiding) henduo   ren   (??yiding) xihuan lisi
         surely   very-many people surely   like   L.
         'Surely many people like Lisi.'

The word order differences can be naturally derived from (8-9). Let's first consider the derivations for (42a) (some unrelated features have been omitted for ease of exposition):
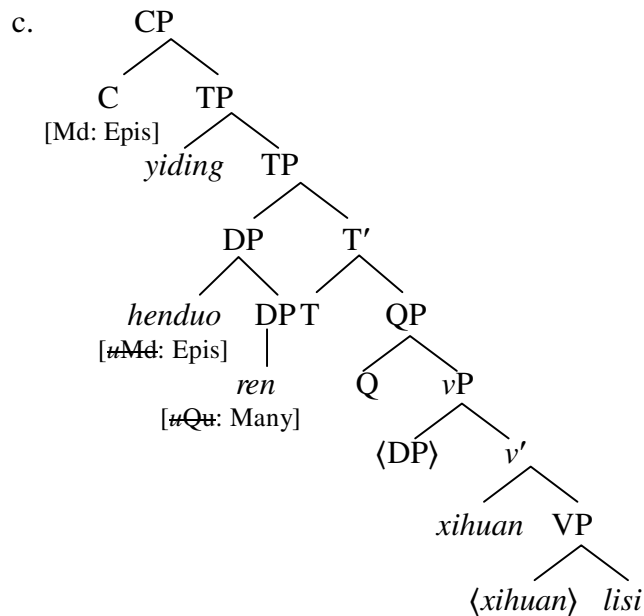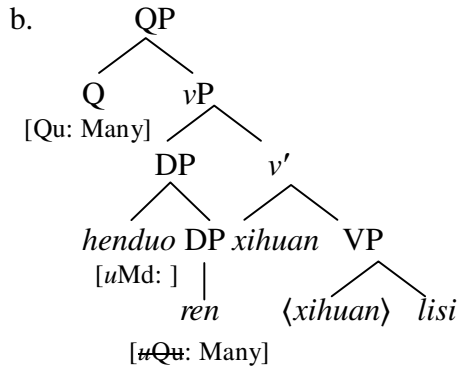
198

(43) a.

```
              vP
             /  \
     zhangsan    v'
                /  \
           xihuan   VP
           [uMd: ]  /  \
                ⟨xihuan⟩ lisi
```

b.

```
              CP
             /  \
            C    TP
      [Md: Epis] / \
          zhangsan  T'
                   /  \
                  T    vP
                      /  \
                 yiding   vP
                         /  \
                ⟨zhangsan⟩   v'
                            /  \
                       xihuan   VP
                      [u̶M̶d̶: Epis] / \
                           ⟨xihuan⟩ lisi
```
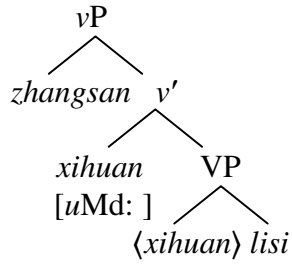
(43a) shows us the stage of derivation that involves *v*P. Here we see that the main verb *xihuan* 'like' is the locus of the *goal*. This is so because *v*P is the focus of the mood expression. The Goal Condition requires the *v* to bear focus, instead of the subject NP. Hence the ensuing pied-piping (selecting *v*P) and delayed-Merge of *yiding* to the *v*P.[38]

Next let's consider the derivations for (42b):

(44) a.

```
          vP
         /  \
       DP    v'
        |   /  \
       ren xihuan VP
       [u̶Q̶u̶: ]   /  \
            ⟨xihuan⟩  lisi
```

---
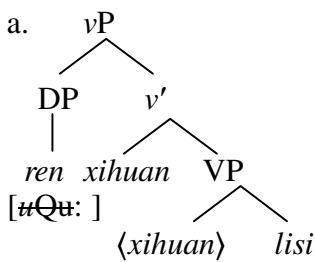
[38] It is still not entirely clear why when the entirely TP is the focus, the *goal* is also the main Aux or the main verb, instead of the subject DP. I will leave the issue aside.

b.

```
              QP
           ╱     ╲
          Q       vP
      [Qu: Many]  ╱  ╲
                DP     v′
              ╱   ╲   ╱  ╲
        henduo  DP xihuan  VP
        [uMd: ]  │        ╱  ╲
                ren  ⟨xihuan⟩  lisi
              [u̶Q̶u̶: Many]
```

c.

```
            CP
          ╱    ╲
         C      TP
     [Md: Epis] ╱ ╲
              yiding  TP
                    ╱   ╲
                  DP      T′
                ╱   ╲    ╱  ╲
          henduo  DP T   QP
          [u̶M̶d̶: Epis]│  ╱  ╲
                  ren  Q    vP
               [u̶Q̶u̶: Many]  ╱  ╲
                          ⟨DP⟩    v′
                                ╱   ╲
                            xihuan    VP
                                    ╱   ╲
                              ⟨xihuan⟩  lisi
```

(44a) shows us the main verb *xihuan* is not the locus of the *goal*. This is so because the focus of the mood expression this time is within the subject DP. Following the semantic fact that the QNP takes scope, I propose that the quantificational element *henduo* 'many' is actually the *realize* of Agree between the *probe* at Q and the *goal ren* 'person', as shown in (44b). Crucially, it is the quantificational element *henduo* that is the focus of the mood expression and the locus of the *goal* for Agree that triggers the merger of the sentence adverb *yiding*. This Agree operation, as shown in (44c), involves selection of the TP for pied-piping. This is allowed by our proposal in (17e): TP-pied-piping is allowed as long as the lexical specification of the FA allows it.[39] Thus Generalization C1 is once again observed: the focus is c-commanded by the FA.

Recall from §3.2.2.2 a theory that only allows SAs to be generated in fixed scope-related positions is unable to account for these facts, because such a theory would either have to force

---

[39] Alternatively, the subject DP is selected for pied-piping in (42b). Either analysis is compatible with our theory.

the subject to topicalize in sentences like (42a), or stipulate that the SAs in (42a) and (42b) are homonyms that have different semantic scopes. As we have already seen, the topicalization approach would wrongly predict that an SA can occur either before or after a definite subject DP. The homonym approach would predict that SAs cannot occur in a post-subject position, since their scope is wider than TP, again this prediction is not borne out.

Let's now look at cases that involve island effects, starting with the following examples from §3.1.2.2.2:

(45) [A: At yesterday's party, there were two strangers. One man talked to Mary. The other man talked to Bill. But I think John knows both of them. B: No….]
    a. John knows only [the man who talked to MARY].
    b. *John knows the man who talked to only [MARY].

According to our discussion in chapter 3, the whole complex DP is in fact the focus associated with the FA *only*. The ill-formedness of (45b) is due to Generalization C1, since *only* doesn't c-command its focus. In our theory, (45a) can again be derived from (7-9). The determiner *the* is the locus of the *goal* [uId] because it is part of the focus.[40] The [uId] selects the complex DP for pied-piping according to (17).

A consequence of this analysis is that general island effects, as well as property (24p) (the CED effect), can be similarly accounted for. Consider the following sentences:

(46) a. *Who did you get jealous because I spoke to?
    b. Who got jealous because I spoke to whom?
    c. %Because I spoke to whom did you get jealous?
(47) a. *What color hair did you meet students with?
    b. Who met students with what color hair?
    c. %Students with what color hair did you meet?
(48) a. *It is John I got jealous because you spoke to.
    b. It is because you spoke to John that I got jealous.
(49) a. *It is red color hair that I met students with.
    b. It is students with red color hair that I met.

The classic account of the ungrammaticality of (a) sentences in (46)-(49) is that they involve moving out of adjuncts (Huang 1982, Fiengo et al. 1988, etc). The grammaticality of (46b) and

---

[40] I assume that all the expressions in the complex DP bear a focal feature, but only one word bears the [uId] feature. See Irurtzun (2007) for a formal treatment of assigning focal features to multiple tokens in the Numeration.

(47b) is account for by the availability of QR of the whole island, whose existence ameliorates movement of the *wh*-phrase due to the fact adjunction cancels the effect of barriers in Chomsky's (1986) barrier framework. This approach, however, is problematic since it is not clear how it can be incorporated in current minimalist framework and it presupposes that pied-piping is not an option in either covert or overt syntax, which runs afoul of examples like (46c), (47c), (48b), and (49b).[41] If we extend our pied-piping analysis for association-with-focus-phenomenon above to wh-movements, we immediately have a natural account. I propose wh-movements and cleft constructions are subject to the following derivations:

(50) *An Agree analysis of wh-movement/cleft construction*

    a. (i) C   X   → C   X

           valued [iF] unvalued [uF]   valued [$i$wh] valued [uF][42]

      (ii) Condition: The head bearing the goal bears the focus of the probe.

    b. [uF]' selects a phrase P(uF) as a candidate for pied-piping.

    c. Merge P(uF) to the edge of CP.

(50a i) is the familiar Agree operation that involves a probe and a goal. (50a ii) comes from the assumption that the probe involved in wh-movement and cleft constructions is a focus-sensitive operator and the condition (8a ii). According to this condition, an expression cannot bear a [uF] feature if it does not bear the focus of the probe. According to the definition of focus given in chapter 3, *because I spoke to whom* and *students with what color hair*, *because I spoke to John*, and *students with red color hair* are the foci of examples in (46), (47), (48), and (49), respectively. The intuition can be made concrete by the following paraphrases of (46c), (47c), (48b), and (49b):

(51) a. Among the alternatives, BECAUSE I SPOKE TO JOHN, BECAUSE I SPOKE TO MARY, BECAUSE I SPOKE TO JILL, etc, which was the cause of your jealousy?

    b. Among the alternatives, STUDENTS WITH BROWN COLOR HAIR, STUDENTS WITH

---

[41] Fiengo et al. (1988: 87) did mention cases of overt pied-piping. However, they only discussed cases of embedded questions constructed from (46) and (47), which are degraded, and curiously did not cover cases of direct questions, which are fully acceptable to some (but not all) my informants. It's not clear to me why direct and indirect questions have this contrast, but the fact remains that overt pied-piping is a viable option at least to some speakers and should not be dismissed out of hand. For speakers that do not allow overt pied-piping in these cases, I assume their grammar has a PF parameter that puts stricter constraints on this option.

[42] The exact workings of wh-movement are not completely made clear in the Agree framework. According to Chomsky (2000), C have (i) a P-feature of the peripheral system (which is also the EPP feature), (ii) an uninterpretable [Q] feature, (iii) a [wh] feature. How they work exactly is not completely explained. Here I arbitrarily assume the relevant feature is [$i$wh] on C and [$u$wh] on the goal, which ensures the assignment of an [EPP] feature to the probe to trigger movement.

BLACK COLOR HAIR, STUDENTS WITH RED COLOR HAIR, etc, which one did you meet?

    c. Among the alternatives, BECAUSE I SPOKE TO JOHN, BECAUSE I SPOKE TO MARY, BECAUSE I SPOKE TO JILL, etc, it is the first one that was the cause of my jealousy.

    d. Among the alternatives, STUDENTS WITH BROWN COLOR HAIR, STUDENTS WITH BLACK COLOR HAIR, STUDENTS WITH RED COLOR HAIR, etc, it is the third one that I met.

Since it is the whole 'island' phrases that are the foci, any of them can in principle bear the [$u$wh] or some A′-related feature. Locality and economy principles ensure that the head that is closest to the C bears the [$u$wh] or some A′-related feature. Thus, the head that bears [$u$wh] is *because* in (46c) and the *n* head in (47c). The next step, (50b), then selects the whole phrase headed by the *goal* for pied-piping (according to (17)), and then in (50c) the pied-piped phrase is merged to the edge of CP. These derivations are very similar to the ones that derive (45a).[43]

A consequence of this approach is that extraction from an island is allowed, as long as the island is not the focus of the focus-sensitive operator that triggers extraction. Examples of this kind do exist, as illustrated below (data from Fiengo et al. 1988):

(52) a. *Who do you think that pictures of are on sale?[44]

    b. ?Who do you wonder which pictures of are on sale?

In (52b), *who* can be extracted from a subject island, whereas in (52a) it cannot. This contrast follows naturally from our analyses of island effects above when we examine the foci that are involved in these sentences. Specifically, it can be shown that in (52a), the focus is *picture of whom* and in (52b), the focus is *who*. There is both syntactic and semantic evidence for this analysis. First, as we have seen in (50) and (51), an interrogative operator is associated with a wh-phrase. The embedded clause in (52b) is headed by such an operator because it is selected by the verb *wonder*. Therefore, the *wh*-phrase *which pictures of whom* is the focus of the interrogative operator. For interpretational reasons, the interrogative operator at the matrix clause naturally cannot associate with the same focus again, but has to associate with *who*. More specifically, (52b) can be paraphrased as follows:

---

[43] There are still many unsettled issues, such as why the sentence *Who do you believe that John met?* is fine whereas *That John met whom do you believe?* is degraded. Also, for speakers that accept the second sentence, it gets an echo-question interpretation. I believe these questions can be answered when we understand more about that semantic nature of focus, especially how the alternatives are chosen in a given sentence.

[44] Again there are important issues I will leave unresolved here, such as the contrast between the ill-formed *Who do you think that pictures of are on sale?* versus the well-formed *Who do you think that the store sells pictures of?* This is known as the Subject-island effect. The contrast is presumably due to information structure considerations.

(53) Among the alternatives JOHN, MARY, BILL, etc, who do you wonder, among the alternatives HIS BABY PICTURES, HIS HIGH SCHOOL PICTURES, HIS COLLEGE PICTURES, etc, which ones are on sale?

There is no such paraphrase possible for (52a), because there is only one focus-sensitive interrogative operator. For this reason, *pictures of whom* is the focus. It can be paraphrased as follows:

(54) Among the alternatives PICTURES OF JOHN, PICTURES OF MARY, PICTURES OF BILL, etc, etc, which ones do you think are on sale?

Second, their different overt pied-piping possibilities confirm the above semantic intuitions:

(55) a. %Pictures of whom do you think are on sale?
     b. *Which pictures of whom do you wonder are on sale?

The phrase *pictures of whom* can undergo movement in (55a), but *which pictures of whom* cannot in (55b). This contrast follows naturally from our analysis above since *pictures of whom* is the focus of the matrix-clause interrogative operator, whereas *which pictures of whom* is not. In previous analyses, the contrast in (52) was accounted for under Chomsky's (1986) *Barrier* framework (see especially Fiengo et al. 1988), which in effect has to stipulate that island effects are obviated when the island is first moved to an A′ position. This analysis does not fit with the current minimalist framework and fails to capture the facts in (53)-(55).[45]

Let's now return to the theoretical issue. Under our approach to wh-movement, the island effects no longer needs to be stipulated, and can be reduced to a general principle of economy and locality. Let's now spell-out the relevant condition:

(56) Shortest Agree: Among possible derivations, the derivation with the shortest distance between the *probe* and the *goal* is preferred than other derivations.

The contrast between (46a) and (46c) can now be accounted as follows. In these sentences, the head *because* is the optimal bearer of [*u*wh] since it bears the focus of [*i*wh] and is closer to C than the other heads that bear the focus of [*i*wh]. The expression *who(m)*, on the other hand, is

---

[45] Another example of extraction from island is Rullmann's (1997) example we discussed in chapter 3: *They hired no linguists who had even read Syntactic Structures*. Here the DP *Syntactic Structures* can be extracted because the complex DP is not the focus of the FA *even*.

not the optimal bearer of [*u*wh] since it is not the closest potential bearer of the feature. As a result, the expression *because* is the optimal head that can bear the [*u*wh] feature, and this allows the consequent selection of the phrase *because I spoke to whom* for pied-piping and internal-Merge. All the other cases of island effects can be accounted for in the same way.

In sum, our Agree analysis of FAs accounts for facts that concern the c-command condition. What is traditionally called island effects is also accounted for. The focus of the probe is in fact not just some expression within the island, but the whole island itself. As a consequence, it is the head of the island phrase that bears the goal [*u*F] feature, which then undergoes Agree with the probe. FA then undergoes delayed-Merge with the phrase P(*u*F). Our account has the advantage of providing a principled account of pied-piping phenomena, whose syntactic effects are observed both in the placement of FAs and the constituents that undergoes wh-movement, and whose interpretational effects conforms the speakers' intuition. Traditional islandhood accounts will have to deny the existence of pied-piping and the relevance of focus-sensitivity in syntax, which would simply run afoul of empirical facts observed here.

4.3.4.5 Generalization C2 – Adjacency

(57) When an FA c-commands its focus, they cannot be separated by a constituent that is not part of the focus, unless other factors intervene.

As we have seen in §3.1.2.2 and §3.2.2.5, there is a general locality condition that requires an FA to be as close to its focus as possible. Generally, in cases where focus is the VP or a constituent within it, when FAs are attached to higher clausal projections or are separated from the focus by non-focus elements, the sentences are unacceptable. These facts can now simply be accounted for by condition (8a ii), which presupposes the Agree operation between the probe and its focus. Let's again consider the following examples from Chinese (repeated from chapter 3):

(58) a. zhangsan zuotian zai tushuguan **jingran** <u>du-le shi-ben shu</u>
  b. *zhangsan zuotian **jingran** zai tushuguan <u>du-le shi-ben shu</u>
  c. *zhangsan **jingran**   zuotian   zai tushuguan <u>du-le   shi-ben shu</u>
    Z.   surprisingly yesterday at   library   read-Pft ten-Cl book
    'I can't believe Zhangsan <u>read ten books</u> at the library yesterday. (He is usually a slow reader.)'
(59) a. zhangsan **jingran** <u>zai baitian</u> shuijiao
  b. *zhangsan <u>zai baitian</u> **jingran**   shuijiao
    Z.   at day   surprisingly sleep

205

'I can't believe Zhangsan sleeps during daytime. (People usually sleep at night.)'

In (58a), the VP constituent *du-le shi-ben shu* is the focus of the operator with [*i*Md] feature. The verb *du* bears the [*u*Md] feature, due to the Goal Condition (8a ii).[46] The feature then selects VP for pied-piping, allowing delayed-Merge of *jingran* to apply. In (58b), the SA *jingran* is not merged to its focus VP, but instead to a higher clausal projection (presumably a TP-internal TopP) that contains the locative adjunct *zai tushuguan*. Since the head of this project is neither the Aux nor the main verb of the sentence, nor does it bear the focus of the probe [*i*Md], it cannot bear the [*u*Md] feature, according to (8a ii). There is then no way for the SA to merge to the projection of this Top head. Same reason rules out (58c). In (59a), on the other hand, the temporal adjunct *zai baitian* is the focus of the operator with [*i*Md] feature. The preposition *zai* therefore is able to bear the [*u*Md] feature, which allows it to select the phrase *zai baitian* for pied-piping and merging the SA to this phrase. In (59b), the SA is merged to the VP, which doesn't contain the focus and therefore the *goal*, so the sentence is unacceptable.

We can conclude the adjacency condition, which suggests syntactic dependency between an FA and the focus, can be derived from our Agree analysis of FAs.

4.3.4.6 Generalization C3 – Clausemate Condition 1

(60) When an FA doesn't c-command its focus, the focus is in the same minimal clause as the FA before it undergoes further independently-motivated A'-movements.

This generalization describes the fact that although some FAs, such as *even* and *also*, do not need to c-command their foci, they have to be somewhat local to the latter. The relevant facts are repeated below:

(61) a. *[John **even** went home] [although he hadn't met his <u>advisor</u>].
    b. *<u>Mary</u> thought [John'd **even** play cello].

These facts can easily be derived from the economy/locality condition (56). In those examples, based on our definition of scope, the position of the focus indicates the scope position is at the edge of the *although*-clause in (61a) and the edge of the matrix clause in (61b), respectively, marked by P below:

---

[46] Or, alternatively, the main verb *du* is assigned [*u*Md] due to its identity as the main verb. Either derivation works for our purpose.

(62) a. *[John **even** went home] [although P he hadn't met his <u>advisor</u>].

    b. *P <u>Mary</u> thought [John'd **even** play cello].

With such scope positions, the closest target for Agree can only be *hadn't* or *met* in (61a) and *thought* in (61b).

4.3.4.7 Generalization C4 – One-to-many association

(63) An FA may have more than one focus.

As we have seen in §3.1.2.2 and §3.2.2.6, an FA can associate with multiple foci. The multiple foci may or may not have different focus-related interpretational effects, but they clearly all invoke the presence of alternatives in interpretation. Under our analysis, those foci will all be potential bearers of an [uF] feature. This understanding allows us to derive the syntax of the following sentences, repeated from chapter 2:

(64) A: zhangsan jintian zhaodao-le gongzuo. lisi zuotian zhaodao-le gongzuo.

      'Zhangsan got a job today. Lisi got a job yesterday.'

    B1: (bu.) ??<u>lisi jintian</u>   **ye$_a$** zhaodao gongzuo.

    B2: (bu.) <u>lisi</u>   **ye$_a$** <u>zai jintian</u>   zhaodao gongzuo.

       no L.   also at today   find   job

       'No. <u>Lisi</u> also got a job <u>today</u>.'

In these examples, both *lisi* and (*zai*) *jintian* are the foci of the probe [iId: AI (additive identification)] feature. They have somewhat different interpretational effects, as we have discussed in §3.1.2.2. Let's call the former the primary focus, the latter the secondary focus. A peculiar property of *ye$_a$* (and the English FA *also*) that distinguishes it from many other FAs is that it is not syntactically sensitive to where the primary focus is in the sentence, as long as they are both in the same minimal clause. In (64B2), for example, *ye$_a$* is not attached to *lisi*. On the other hand, the contrast between (64B1) and (64B2) shows that it is the *secondary* focus that plays the role of being the locus of the [uId] feature. Further examples can be found where the (quantified) subject DP bears the secondary focus:

(65) [zai taiwan, hen shao ren you lanbaojini]

    'In Taiwan, very few people have Lamborghinis.'

    a. ye$_a$ <u>hen shao ren</u>   you <u>siren   feiji</u>

b. *<u>hen shao ren</u>   ye$_a$ you   <u>siren   feiji</u>
   very few people also have private airplane
   'Also very few people have private jets.'

The above examples again show the quantified subject DP as the secondary focus has to be one of the loci of the [uId] feature. To account for the syntax of *ye$_a$*, we can add the following lexical specification to *ye$_a$*:

(66) Merge site of *ye$_a$*: The secondary focus is the target of delayed-Merge, unless the Aux or the main verb is closer to the probe.

As for the primary foci *lisi* and *siren feiji* in (64) and (65), we have seen they also have syntactic dependencies with the probe, since the former have to be in the *scope* of the latter (according to the definition of scope). I take this as the evidence they also bear an [uId] feature and enter Agree with the probe, establishing multiple-Agree between one probe and several goals. It is this Agree relationship that limits the syntactic distributions of the primary focus. However, the goal [uId] is distinct from the [uId] on the expression bearing the secondary focus in that they do not trigger pied-piping and delayed-Merge of the FA. I will not try to provide a full analysis of the multiple-Agree constructions here (which I suspect will be similar to analyses of other multiple-Agree constructions such as multiple-wh-movement and those constructions discussed in Hiraiwa (2005), Boeckx (2008), etc), since that would go beyond the core issues of sentence adverbs. Here I'd to simply point out the following general condition for sentences with a focus-sensitive operator associated with multiple foci:

(67) Focus condition: In a multiple-foci construction, if expression X bears one of the foci of an expression Y with a probe [F], X bears the goal [uF] of Y.[47]

In sum, with the addition of the lexical specification of how Agree applies for the syntax of *ye$_a$*, we have a principled account of the syntax of multiple-foci. It is difficult to imagine how analyses that disregard syntactic dependencies between a focus-sensitive operator and its foci can deal with relevant facts.

---

[47] This is sort of a mirror-image of (8a ii). However, (8a ii) is concerned with the syntax of FAs, this condition is not.

4.3.4.8 Generalization D1– ECP/Island effects

(68) Island/ECP effects: When an FA is attached to a DP or PP, the constituent [FA DP/PP] is c-commanded by its scope position and is subject to locality constraints and displays ECP effects with respect to that position.

We have seen in chapter 3 that when an FA is attached to a non-clausal constituent such as DP and PP, the resultant [FA DP/PP] constituent has to be 'local' to its scope position and obeys some ECP-like principles. More specifically, we have shown there are four sets of facts: (i) subject-object asymmetry, (ii) indicative-subjunctive asymmetry, (iii) island effects. Now these facts have natural accounts under our Agree analyses. Let's consider them one by one. First, let's consider cases of subject-object asymmetry:

(69) a. ⌜John has requested ⌜Bill study only <u>physics</u>⌝⌝.
    b. John has requested ⌜only <u>Bill</u> study it⌝.

In our analysis, the wide-scope reading of *only physics* in (69a) is derived from Agree with a probe at a functional head X and the goal at *physics*. Delayed-Merge then merges *only* to *physics*. Next, an EPP feature at X triggers internal Merge of *only physics* covertly to the edge of XP. The consequence of this analysis is that the subject-object asymmetry in (69) can then be assimilated to well-known (but still ill-understood) cases of wh-movement where object extraction is acceptable but subject extraction is barred:[48]

(70) a. ?What do you wonder when John bought?
    b. *Who do you wonder when bought these books?

A modern minimalist account of the contrast in (70) can be found in Rizzi (2006), who argues the following condition on movement:

(71) *Criterial Freezing*
    A phrase meeting a criterion is frozen in place.

---

[48] Note, however, pied-piping makes these sentences both acceptable:

(i) a. When John bought *what* do you wonder?
    b. When *who* bought these books do you wonder?

Translated into our Agree account of FAs, we can then state the following condition:

(72) An expression α can only have one [EPP] probe.[49]

Condition (72) accounts for the subject-object asymmetries in (69) and (70): subject cannot undergo further movement because it will have to be probed twice. I assume this holds without getting into details of Rizzi's arguments.[50]

    I will leave the indicative-subjunctive asymmetry to the side, since the issue is poorly understood and as far as I know there has been no systematic Agree analysis of relevant asymmetry (but see Szabolcsi 2006 and references cited there for discussions on the role of tense in determining islandhood).

    If our discussions of the Generalization C1 are on the right track, a number of island effects can be accounted for by the economy/locality principle (56). We can also account for the ban on object extraction in certain non-island environments:[51]

(73) a. %⌜I knew that he had learned only <u>Spanish</u>⌝. (I didn't know he had learnt any other language)

    b. ⌜I knew only that he had learned <u>Spanish</u>⌝. (I didn't know he had learnt any other language)

(74) a. youngsu-nun ⌜haksayngtul-ekey <u>lesiamal</u>-man paywu-lako⌝ yokwuha-yss-ta
    Y.-Nom   students-Dat   Russian-only learn-Comp demand-Past-Dec.
    'Youngsu demanded that the students only learn <u>Russian</u>.'

    b. ⌜Y.-nun haksayngtul-ekey <u>lesiamal</u>-lul paywu-lako-man yokwuha-yss-ta⌝[52]
    Y.-Nom students-Dat   Russian-Acc learn-Comp-only demand-Past-Dec.
    'Youngsu demanded only that the students learn <u>Russian</u>.'

(73a) shows an [FA DP] constituent with DP as the focus in an embedded clause headed by an overt complementizer *that* cannot take scope over the matrix clause for some speakers. The wide scope reading, however, can be expressed by attaching the FA to the whole embedded CP, as shown in (73b). Similarly situations are found in Korean. (74a) shows that an [FA DP] constituent in an embedded clause headed by an overt complementizer *-lako* cannot take the matrix-clause scope. On the other hand, the matrix-scope reading of FA can be expressed by

---

[49] In Chomsky (2008), a subject wh-phrase is able to undergo two independent movements in parallel. This analysis cannot capture the contrast in (70). See also Bošković (2009) for some discussion.
[50] Alternatively, the asymmetry can be derived from information structure considerations. See note 44.
[51] See also note 30 in chapter 3.
[52] This sentence is in fact ambiguous between wide-scope and narrow-scope readings for some speakers.

attahching FA to the whole embedded CP, as shown in (74b). Syntactic theories that accounts for restrictions on syntactic dependencies solely on island conditions will have a hard time accounting for this contrast. In our analysis, on the other hand, the contrasts above follow from condition (56). The complementizer is the closer potential bearers of [*u*Id]. This is so presumably because the whole embedded CP can be regarded as a secondary focus. Thus, in our analysis, (73b) and (74b) actually have the following focus structures, respectively:

(75) a. 「I knew only <u>that he had learned Spanish</u>」. (I didn't know he had learnt any other language)

   b. 「Y.-nun <u>haksayngtul-ekey lesiamal-lul paywu-lako</u>-man yokwuha-yss-ta」
     Y.-Nom students-Dat   Russian-Acc learn-Comp-only demand-Past-Dec.
     'Youngsu demanded only <u>that the students learn Russian</u>.'

In sum, the various island and ECP effects in sentences with [FA DP] constituent again follow naturally from our Agree analysis of FAs and some independently motivated principles on syntactic dependencies, such as Shortest Agree (56) and the one that derives the ECP effect (72).

4.3.4.9 Generalization D2 – Clausemate Condition 2

(76) When an FA is attached to a clausal projection, its scope is the minimal clause that contains the FA.

As mentioned before, this condition deals with examples like the following:

(77) a. I knew 「he had only learnt <u>Spanish</u>」. (I knew he hadn't learnt any other language.)
   b. % 「I knew he had only learnt <u>Spanish</u>」. (I didn't know he had learnt any other language.)
   c. 「They were only advised to learn <u>Spanish</u>」. (They were not advised to learn any other language.)

These examples are readily accounted for under our Agree analysis of FAs, although some parametric settings must be involved to account for variations among speakers.

    The lack of wide scope readings in (77a) for some speakers can be readily accounted for by the economy/locality condition (56). Since there is a closer potential bearer of the goal, namely the verb of the main clause *knew*, assignment of [*u*F] to the farther head *had* is barred. For speakers who accept (77b), on the other hand, the economy/locality condition (56) has to be relativized or loosened. Again there are several possible analyses for this parametrization. One is

that the condition simply doesn't need to hold for some speakers. For those speakers, either the matrix verb or the embedded Aux can bear the [uF] feature, simply because condition (56) doesn't hold in here. As a result, 'long-distance' Agree is allowed. Evidence for such approach comes from the existence of long-distance agreement that involves φ-features, as discussed in Spenser (1991), Miller (1992), Corbett (2008), and various other sources. For long-distance Agree to work in these cases, the condition has to be loosened. The other possibility is to let the locality condition be relativized to whether the potential goal bears focus or not:

(78) Shortest Agree－RELATIVIZED VERSION: If the *goal* bears the focus of the *probe*, then among possible derivations, the derivation with the shortest distance between the *probe* and the *goal* is preferred than other derivations.

In this looser version of locality condition, *knew* is no longer an intervener in (77b) since it does not bear focus of the probe. The [uId] feature is assigned to *learnt* without violating (78). Note this condition may also account for the existence of sentences such as *John likes only Mary* and examples in (69). I will not choose among these approaches, which will warrant a separate topic of research, but the point should be clear that an Agree analysis is on the right track.[53]

As for the lack of embedded-clause scope reading for *only* in (77c), it follows from the specification of directionality parameter in (8a iii): the controller has to c-command the target.[54] An embedded-clause scope reading for *only* would place the functional head bearing [iId] feature at the edge of the embedded clause, which cannot c-command the Aux *were* of the matrix clause. The latter therefore cannot be the target of Agree.

In any event, our analysis offers a principled account for the Clausemate Condition 2, which is about cases where FAs are attached to clausal constituents, based on independently-motivated syntactic principles and specifications of parametric settings about locality of syntactic dependency.

4.3.4.10 Generalization D3 – Intervention Condition

(79) When an FA is attached to a clausal projection, it can not be intervened by more than one verbal head from its scope position.

Similarly to Generalization D2, this generalization can be readily accounted for in our analysis by the economy/locality condition (56). Let's review some relevant examples:

---

[53] Optional violations of locality are in fact quite robust with respect to FAs. See note 14 and note 19 in chapter 3.
[54] See also Baker (2008: ch 5) for some general discussions of this parameter.

(80) a. ⌜John could **only** have been dating <u>Mary</u>⌝. (He couldn't have been dating others.)

    b. %⌜John could have **only** been dating <u>Mary</u>⌝. (ditto)

    c. %⌜John could have been **only** dating <u>Mary</u>⌝. (ditto)

In all of the above examples, *only* has the wide scope, and it is shown it cannot appear after the second auxiliary verb at least for some English speakers. Under the Agree analysis, this follows from (56), which entails no potential bearers of [*u*F] can intervene between the probe and the goal. In (80a), the bearer of the [*u*F] feature is the highest auxiliary verb *could*, so the FA *only* is right-adjoined to it. In (80b) and (80c), the goals are presumably the second or the third auxiliaries. They are out since the locality condition is violated. For speakers that allow the wide-scope reading of (80b,c), condition (56) simply doesn't need to hold, since it is a parameter, or that they adopt a looser version of (56), a possibility which is motivated by other cross-linguistic cases of 'long-distance' agreement phenomena discussed above.

4.3.4.11 Generalization D4 – Overt movement and blocking

(81) If overt QR is available or scrambling has a semantic effect in a language, an [FA DP/PP] constituent cannot occur in a position that doesn't mark its scope in overt syntax.

In §3.1.2.1, we saw languages such as Chinese and German allow overt focus movement/ scrambling, whereas English adopts the option of covert focus movement. The contrasts between these languages not only show that there is a parameter that determines whether a language adopts overt focus-related movements or covert ones, but also further provides support for our analyses in (8d) and (9d). Let's review the relevant examples:

(82) a. *lisi yaoqiu xuesheng yanjou [zhiyou <u>yazhou de yuyan]</u>

      L. request student   study   only Asia   DE language

    b. lisi yaoqiu ⌜xuesheng [zhiyou   <u>yazhou de yuyan]</u>ᵢ cai yanjou tᵢ⌝

      L. request student   only   Asia   DE language CAI study

      'Lisi requested that students only study Asian languages.'

    c. ⌜lisi [zhiyou <u>yazhou de yuyan]</u>ᵢ cai yaoqiu xuesheng yanjou tᵢ⌝

      L. only   Asia   DE language CAI request student   study

      'It is only <u>Asian languages</u> Lisi requested that students study.'

d. ⌜[zhiyou yazhou de yuyan]ᵢ lisi cai yaoqiu xuesheng yanjou tᵢ⌝

only   Asia   DE language L. CAI request student   study

'It is only <u>Asian languages</u> Lisi requested that students study.'

(83) a. Die Studenten in der DDR wurden gezwungen ⌜nur <u>Russisch</u> zu lernen⌝

the students   in the GDR were   required   only Russian to learn

'The students in the GDR were required to only learn <u>Russian</u>.'

b. ⌜Die Studenten in der DDR wurden [nur <u>Russisch</u>] gezwungen tᵢ zu lernen⌝

the students   in the GDR were   only Russian   required   to learn

'It is only Russian the students in the GDR were required to learn.'

c. ⌜[Nur <u>Russisch</u>] wurden die Studenten in der DDR gezwungen tᵢ zu lernen⌝

only Russian   were   the students   in the GDR required   to learn

'It is only Russian the students in the GDR were required to learn.'

d. %...weil ⌜die Studenten in der DDR [nur <u>Russisch</u>] gezwungen tᵢ wurden zu lernen⌝

since the students   in the GDR only Russian   required   were   to learn

'…Since it is only Russian the students in the GDR were required to learn.'

The Chinese examples in (82) and German examples in (83) show that an [FA DP] constituent needs to undergo focus movement overtly to a scope-related position in order for the sentence to be well-formed. In our Agree analysis of FAs, these sentences could receive a principled account if we follow the standard theory of movement according to which an agreeing probe can trigger the assignment of an EPP feature to the probe (which is also termed a *P-feature* in Chomsky (2000: 108, 144)). This EPP feature can trigger either covert or overt movement of the [FA DP] constituent to the edge of XP, X being the functional head bearing the [Id] feature (our derivation (8d)). Note, however, that the landing sites of [FA DP] in the examples are often not the scope positions, the edge of XP, since the former are often between the subject and the verb, instead before the subject position. In other words, we expect (84a), but what we get is (84b):

(84) a. [$_{XP}$ [FA DP]ᵢ [$_{X'}$ X [$_{TP}$…[$_{vP}$…tᵢ…]]]]

b. [$_{XP}$ X [$_{TP}$ …[$_{YP}$[FA DP]ᵢ [$_{Y'}$ Y [$_{vP}$…tᵢ…]]]]]

We can account for this discrepancy by allowing the result of internal Merge to be spell-out not at the edge of the highest agreeing probe, but at the edge of the *v*P that contains the "intermediate" probe bearing purely formal EPP features.[55] Similar analyses have been proposed for cases partial *wh*-movement constructions in languages like Bahasa Indonesia in Groat and

---

[55] This kind of EPP feature is also called *edge feature* (Chomsky 2005, 2008).

O′Neil (1996), Fanselow and Ćavar (2000), and Fanselow (2006). Under such a view, (84b) in fact has the following structure:

(85) $[_{XP} \text{[FA DP]}_i [_{X'} X [_{TP} \ldots [_{vP}\text{[FA DP]}_i [_{v'} v [_{vP}\ldots t_i \ldots]]]]]]$

I assume such an analysis is on the right track, and leave the details for future research.[56]

4.3.4.12 Generalization E – Surface Effect

(86) An FA in a derived syntactic position has interpretational effects.

As we observed in the previous chapter, an FA in a shifted position has interpretational effects. The relevant examples are repeated as follows:

(87) a. [You said that John often buys novels and textbooks, but I think…]
John buys [only$_1$ novels] often$_2$.
b. [You said that John often buys novels and textbooks, but I think…]
*John often$_2$ buys [only$_1$ novels].
(88) a. *Often$_2$, [probably$_1$ only John] buys beer.
b. [Probably$_1$ only John] often$_2$ buys beer.

This generalization in fact is not limited to just focusing adverbs, but applies to adverbial adjuncts in general as well as quantified DPs, as we have shown in chapter 3[57]. In general, these facts can be regarded as subcases of topicalization, since the movements are optional (without topicalization the expressions can still get wide scope), and there is no prosodic stress associated with the moved element. Since topicalization is a separate process from the derivations of focusing adverbs and sentence adverbs, and it is a research topic of its own, I will not attempt to provide a detail analysis here, but simply point out some possible directions.

---

[56] There are two families of influential alternative theories that deal with TP-internal topic and focus, which don't seem to be able to account for the data discussed here. These approaches are (i) discourse-template approach: specific discourse-related interpretations are associated with peripheral positions (Chomsky 1995, 2000, 2001, Neeleman and Van de Koot 2008, Tsai 2008b); (ii) expanded/exploded vP approach: languages may have an expanded vP edge which mirrors the split CP domain first proposed by Rizzi (1997) (Jayaseelan 2001, Belletti 2004, Grewendorf 2005, Aldridge 2010). The first approach is unable to generate any inflectional marking on the goal. The second approach is not able to account for the fact that focus operators can have wide scope over tense and the fact that sentence adverbs have C^0 properties. It is still possible all these approaches can be integrated somehow, which I will leave to the side.

[57] Recall that weak island effects that involved topicalized adverbs are also part of the same phenomena (property 24q) (see note 52 of chapter 3).

There are generally two theoretical approaches to topicalization in the minimalist framework, the EF approach and the P-feature approach. According to the first approach (Chomsky 2001, 2008), there is no such formal feature [topic] at a functional head in the numeration. Instead, topicalization is triggered by an optional EPP/OCC feature (aka edge feature EF) of a phase head PH, which can seek any AP/DP in the phase and raise it to Spec-PH. The result of the movement has interpretational effects, which in our cases forces the moved AP/DP to receive wide-scope interpretation. According to the second approach (Aboh 2010), [topic] is a formal feature that enters the numeration just like other formal features and triggers Agree and Move just like other operations do. The result of movement in our cases corresponds to the wide scope interpretation, because the surface syntax feeds the semantic interpretation directly. Either approach requires deeper understanding of the nature of the relationship between syntax and semantics especially in light of recent developments of phase-based syntactic theories, where NS, Φ, and Σ proceed in parallel, which is still a mostly uncharted territory.

A further important issue is that in our analysis the moved elements here are elements that undergo delayed-Merge, instead of set-Merge. The fact that movement of these expressions still display scope-related effects in spite of the special way they enter syntactic derivations also need to explored in more detail. I will leave this issue for future research.

4.3.4.13 Generalization F – Intervention Condition 2

(89) a. An FA_1 and its scope position cannot be intervened by an [FA_2 DP] if any part of [FA_2 DP] is the focus of FA_1.
   b. α intervenes between β and γ if α c-commands γ and α does not c-command β.

This generalization covers facts that some FAs are forced to topicalize or attach to the subject DP when the subject DP contains an FA that is part of the former's focus. Relevant examples are repeated below:

(90) [I am teaching classes at this university.]
   a. *[$Only_2$ female students] are $usually_1$ very diligent.
   b. $Usually_1$, [$only_2$ female students] are very diligent.
(91) a. *[$Only_2$ John]$_{F_1}$ has $probably_1$ written emails to Mary.
   b. $Probably_1$ [$only_2$ John]$_{F_1}$ has written emails to Mary.
   c. *[No_2body]$_{F_1}$ has $unfortunately_1$ passed the GRE exam.
   d. Unfortunately [no_2body]$_{F_1}$ has passed the GRE exam.

In cases where FA_1 is attached to the [FA_2 DP] constituent (e.g. 91b), the generalization can be translated into some version of locality principle such as (56)[58]. Since *only* is the closest potential bearer of the [uMd] feature, it has to be the locus of the goal. Following the pied-piping option (17a), the subject DP is selected for pipe-piping for delayed-Merge with FA_1. In cases where the FA_1 is not attached to [FA_2 DP] but to the TP instead (e.g. 90b), the result can again be derived from (56) (so vP is not the target for delayed-Merge), but now the pied-piping option (17e) is chosen. Adjunction to TP is chosen over adjunction to DP since idiosyncratic c-s features of *usually* bars its merger with DP constituents.

4.3.4.14 Sentences with multiple FAs and Generalization G – Late insertion effect

(92) When two FAs share the same focus, the FA with narrow scope c-commands the FA with wide scope.

All the cases of sentences with multiple FAs discussed in chapter 3 can now receive straightforward accounts under our Agree analysis, since they conform with the generalizations discuss above, which we have shown can be derived from various Agree-related operations we presented in (7-9). Specifically, the Goal Condition (8a ii) allows the focused expression of a given FA to be chosen as the locus of the *goal*. This allows an FA to attach to various positions in accordance with its information structure. Property (24m) now also follows naturally, adverbs can occur in any order as long as proper syntactic configurations required for Agree in (8) and (9) are satisfied. To see this, let's consider the following examples from chapter 3 again:

(93) [Lisi doesn't know how to sharpen a pencil…]
    a. ta jingran$_1$ changchang$_2$ [yong ya yao   qianbi]$_{F_1/F_2}$
    b. ta changchang$_2$ jingran$_1$ [yong ya   yao   qianbi]$_{F_1/F_2}$
      he often   surprisingly use tooth bite   pencil
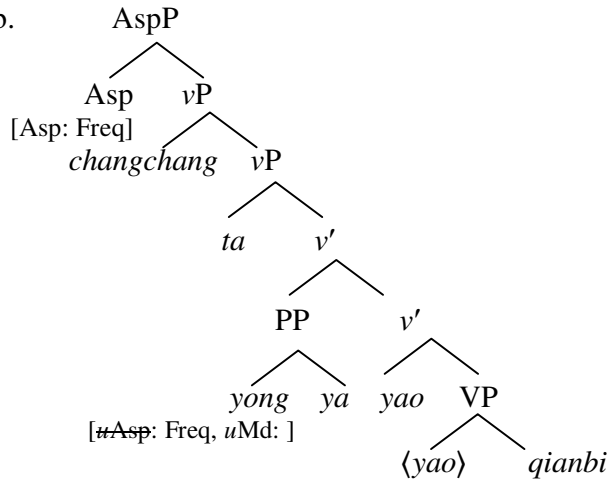   'He often uses his teeth to bite pencils. I can't believe someone can do such a thing.'

Under our Agree analysis, (93b) can be derived as follows:

---

[58] Alternatively, we have multiple-Agree here and attachment to the subject DP is determined by some version of the Intervention Parameter (41).
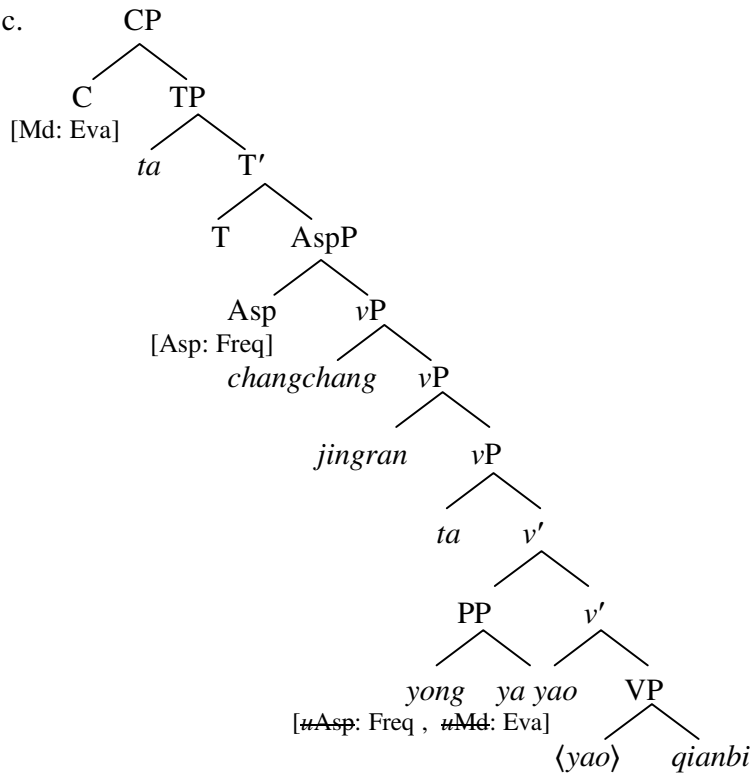
(94) a.

vP
- ta
- v′
  - PP
    - yong ya yao
    - [*u*Asp: , *u*Md:]
  - v′
    - VP
      - ⟨yao⟩ qianbi

b.

AspP
- Asp
  - [Asp: Freq]
- vP
  - changchang
  - vP
    - ta
    - v′
      - PP
        - yong ya yao
        - [~~*u*Asp~~: Freq, *u*Md:]
      - v′
        - VP
          - ⟨yao⟩ qianbi

c.

CP
- C
  - [Md: Eva]
- TP
  - ta
  - T′
    - T
    - AspP
      - Asp
        - [Asp: Freq]
      - vP
        - changchang
        - vP
          - jingran
          - vP
            - ta
            - v′
              - PP
                - yong ya yao
                - [~~*u*Asp~~: Freq , ~~*u*Md~~: Eva]
              - v′
                - VP
                  - ⟨yao⟩ qianbi
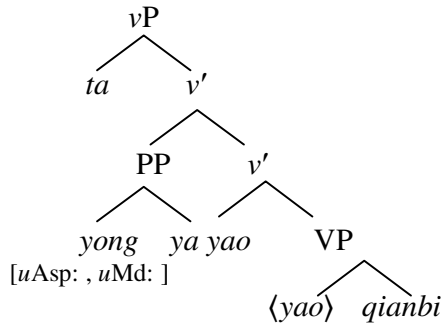
(94a) shows that the preposition *yong* 'use/with' is the head that is the bearer of both [*u*Asp] and [*u*Md] features. (94b) shows the derivations that involve the merger of the frequency adverb *changchang* 'often'. As a focusing adverb, its merger is triggered by Agree between the probe on the Asp head and the *goal* on the P head *yong*, the selection of *v*P for pied-piping, and the consequent merger of the adverb and the *v*P.[59] In (94c), another Agree operation applies due to the merger of the C head. This Agree now happens between the probe on C and the *goal* on P. Again *v*P is selected for pied-piping. Crucially, since *changchang* is not the focus of the probe on C, it doesn't need to be included in the phrase for pied-piping. The evaluative mood adverb *jingran* 'surprisingly' then is merged with the 'smaller' *v*P. The resultant effect is that *jingran* is c-commanded by *changchang* as shown in (94c). Thus Generalization G is again naturally derived from our proposal that grammar allows the option of delayed-Merge, and that the locale of the realization of feature valuation is sensitive to which expression is the focus of the probe (the Goal Condition in (8-9)).

To sum up §4.3.4, reviewing the various generalizations we gathered in chapter under the lens of our Agree analysis of FAs and SAs suggests that when it comes to extremely complex linguistic phenomena, it is still possible to have a successful marriage of comprehensive empirical facts and theoretical coherence, which shows our Agree analysis is on the right track. Now let's examine the other main properties of sentence adverbs.

4.3.5 Sentence adverbs are a heterogeneous group

The fact that sentence adverbs form a heterogeneous group in terms of their morphosyntactic distributions now follows from lexical specifications of individual adverbs in question. More specifically, as the function of FAs and SAs is realization of feature valuation, the delayed-Merge process is also constrained by idiosyncratic lexical properties as to the syntactic categories of the constituents the adverbs can merge with, whether pied-piping is involved, and whether left-adjunction or right-adjunction is the chosen option. The latter two parameters resemble the parameters that govern the morphosyntactic distributions of clitics. Klavans (1985), for example, argues that there are three parameters about clitic placement: (i) it attaches either on the left edge or right edge of the phrase; (ii) it attaches either to the left or right of the element that is on the edge of the aforementioned phrase; (iii) it will be phonologically dependent either to the left or the right. Apparently, any theory of adverbial syntax has to deal with these parameters across different lexical items. Our theory in (8) and (9), being minimalist in spirit, in fact have the advantage of allowing the possibility of these lexical specifications, since

---

[59] Alternatively, the PP is the one selected for pied-piping. Either approach can derive the correct results.

delayed-Merge as a syntactic operation should involve formal feature operations, just like standard cases of set-Merge. Theories that are based on prefab syntactic templates (e.g. Cinque 1999) are not able to account for lexical-specific properties of sentence adverbs. Theories that are solely concerned with semantic and phonological interface issues are also not able to account for the relevant data (e.g. Ernst 2002). A brief sketch of how our theory can account for lexical variations should suffice to show the difference. Consider now the following sentences:[60]

(95) a. Mary probably likes John.

    b. Probably, Mary likes John.

    c. Mary likes probably only John.

    d. Mary is probably gone.

(96) a. Mary sure likes John.

    b. *Sure, Mary likes John. (ok with different meaning)

    c. *Mary likes sure only John.

    d. Mary is sure gone.

(97) a. Mary SO aced that exam.

    b. *SO, Mary aced that exam. (ok with different meaning)

    c. *Mary aced SO only that exam.

    d. Mary is SO gone.

(98) a. ??Mary fortunately likes John. (ok with parenthetical intonation)[61]

    b. Fortunately, Mary likes John.

    c. ??Mary likes fortunately only John. (ok with parenthetical intonation)

    d. ??Mary is fortunately gone. (ok with parenthetical intonation)

(99) a. ??Mary surprisingly likes John. (ok with parenthetical intonation)

    b. Surprisingly, Mary likes John.

    c. ??Mary likes surprisingly only John. (ok with parenthetical intonation)

    d. *Mary is surprisingly gone. (ok with parenthetical intonation)

(100) a. Mary nevertheless likes John.

    b. Nevertheless, Mary likes John.

    c. ??Mary likes nevertheless only John. (ok with parenthetical intonation)

    d. Mary is nevertheless gone.

(101) a. *Mary however likes John. (ok with parenthetical intonation)

    b. However, Mary likes John.

---

[60] All the judgments here come from the same informant.

[61] As mentioned in chapter 1, I will not discuss the parenthetical cases in this thesis. They are presumably licensed by different mechanisms that may or may not be related to focus-sensitivity and Agree. See

c. *Mary likes however only John. (ok with parenthetical intonation)

d. *Mary is however gone. (ok with parenthetical intonation)

The above examples show the some distribution possibilities of eight sentence adverbs in English. It is clear from these examples that these adverbs differ with regard to the category of the constituents they can adjoin to. Under our Agree analysis, these differences can be represented by c-s properties of the adverbs, which constraints their delayed-Merge possibilities. Thus something like the following has to be in the grammar of English speakers:

(102) *C-s specifications of some sentence adverbs in English*

Sure/SO: T, vP

Probably: TP, T, vP, DP

Surprisingly/fortunately/however: TP, (T), (vP), (DP)

Nevertheless: TP, T, vP, (DP)

Note that something similar to (102) may be necessary with regard to the morphological rules that involve inflectional affixes in some languages. Swahili gender markers, for example, are specified as to which syntactic categories of the root they attach to, as shown in the following Swahili gender marking (Welmers 1973: 171):

(103) M-tu   m-moja   a-likuya

SG-person(1/2) 1-one   1-came

'one person came'

In this example, the agreement affix *m-* is attached to numerals, where as *a-* is attached to verbs.[62]

4.3.6 Cross-linguistic variation

Just as different sentence adverbs in a language may have different lexical specifications with regard to their morphosyntactic distributions, sentence adverbs are also known to show consistent cross-linguistic differences, which in our approach can be regarded as different parametric settings with regard to direction and pied-piping of delayed-Merge, whether locality can be violated, and whether covert or overt movement is involved. These sorts of parameters are

---

[62] Recall the existence of focusing affixes discussed in note 10 in chapter 3. Their morphosyntactic distributions are likewise determined by their lexical specifications.

known not to be limited to adverbial syntax, but found in a variety of linguistic phenomena: directionality parameters arguably exist in set-Merge, locality parameters determines whether φ-features can be realized far away for their controllers or not, and covert/overt Move parameter arguable also exist for A′-movements, such as QR. Some of these parameters have been explored in some detail, but some are not. Here I will simply list the parameters that are relevant for the syntax of sentence adverbs, and leave a proper theory of these parameters for the future work.[63]

(104) *Directionality/pied-piping parameter settings of FAs attached to various hosts:*[64]

	Auxiliary verb	lexical verb	VP
Chinese:	left	NA	left
English:	right/(left)	NA	left
Italian[65]:	right/(left)	right (V[+tense])	left (V[-tense])
French:	right	right (V[+tense])	left (V[-tense])

(105) *Locality parameter settings:*

	Local Agree	Non-local Agree
Chinese:	✓	
English1:	✓	
English2:	✓	✓
French[66]:	✓	✓

(106) *Overt vs. covert Move parameter settings of [FA DP] constituents*

	Overt Move	Covert Move
Chinese	✓	
English[67]		✓
German	✓	
Korean	(✓) (limited to its relative positions to adverbs)	

---

[63] For example, one may explore whether and how these micro-parameters are correlated with other micro-parameters of the languages at issue, and whether and how these micro-parameters are related to the macro-parameters of those languages (e.g. analyticity parameter, etc). More generally, one may also investigate how these parameters bear on Chomsky's (2007, 2009) recent view that variety of language falls to the lexicon and to the ancillary mappings involved in externalization.

[64] See Williams (1994: 192) for a similar formulation.

[65] Italian and French facts are discussed in Kayne (1989) and Belletti (1990, 1994).

[66] The French facts are discussed in Ernst (2002: 375).

[67] Although English does overt focus movements, such as those in cleft constructions and focus-inversion constructions, they have different functions and are not the major strategy for forming focus.

Note the first two parameters can also be easily found with inflectional morphology. Genitive case markers, for example, are realized as affixes attached to the head nouns in languages like German and Russian, but as a clitic attached to the whole DP in English. Inflectional affixes can either be prefixes or suffixes. Non-local φ-feature agreement is also attested in Spenser (1991), Miller (1992), Corbett (2008), and the cases discussed in note 37.

To sum up §4.3, we have shown with our Agree analysis of focusing adverbs and sentence adverbs, the six major properties of sentence adverbs in (23) that have puzzled previous approaches now can all be subject to principled accounts that fit snugly with minimalist framework. The central idea is that focusing adverbs and sentence adverbs are derived in a way quite similarly to the way inflectional affixes are derived, whose forms and distributions are not determined by inherent properties of the hosts of the adverbs/ affixes, but by the syntax contexts, namely the controllers of the Agree operation, by the presence of a suitable target, and by general principles that govern syntactic dependencies.

4.4 Conclusion

In this chapter, I presented the main proposal of this thesis. I began with introducing state-of-the-art ideas and concepts of the Agree theory in minimalist generative grammar. It is shown that the theory is conceptually desirable in that it is driven by Ockham's Razor, aiming to eliminate as many as unnecessary theoretical constructs as possible, while as the same time it is able to achieve greater generality in that it covers not only facts of syntactic displacement but also syntactic derivations that interacts with inflectional morphology. One major consequence is the possibility for uninterpretable features to enter the numeration unvalued. This allows lexical items and affixes alike to merge 'late' for the sake of feature valuation. The proposal I presented in section 2 cashes in on this consequence and features the operation *delayed-Merge*, which follows from the aforementioned desideratum of Agree theory. Another major feature of the proposal is that focus to mood is what case to agreement (8a ii). Under our analysis, inflectional morphology and the syntax of focusing/sentence adverbs are virtually identical, which is expected in a system where NS, Φ, and Σ proceed in parallel. In section 3 I evaluated the proposal by revisiting the major properties/puzzles of sentence adverbs which couldn't be solved in previous approaches. It was shown that these properties/puzzles can be derived from our proposed Agree theory. Special attention was paid to the workings of the focus-sensitivity of focusing adverbs and sentence adverbs. All the generalizations about focusing adverbs and sentence adverbs we deliberated in chapter 3 received principled accounts that follow naturally from our proposed theory depicted in (7-9) and from general principles that are independently motivated in previous syntactic theories. In the next chapter, I will address further theoretical

consequences and some outstanding issues to be resolved in future research.

5. Conclusion and outlook

In this chapter, we will review what we have achieved in this thesis, explore the morals we can learn from them, and lay out the outlook we have on future studies of relevant topics.

5.1 Overview of what has been achieved

In the beginning of the thesis I have shown that syntactically, sentence adverbs are one of the strangest groups of words in generative grammar, because the current studies are either unable to account for many of their important syntactic distribution facts, or are forced to be limited to describing the recalcitrant facts without overall theoretical coherence. This thesis sets forth to unravel the strangeness of these expressions by first providing a working definition under the current generative framework (chapter 2). It was shown that they do form a syntactic natural class based on the fact that they have properties of adverbial adjuncts and they have various syntactic properties of C^0 elements. These properties are sufficient enough to distinguish sentence adverbs from other expressions, so they are regarded as sentence adverbs' defining properties. In the next leg of our journey (chapter 3) we set forth to unfurl another set of the core 'weird' properties of sentence adverbs, which have generally been sidelined in the literature. These are their focus-sensitivity properties. To show that sentence adverbs do have these properties, we again start from tasking ourselves to provide working definitions to two notions, focus and focus-sensitivity, which play the fundamental role in determining the syntactic distributions of focus-sensitive adverbial adjuncts. Based on these working definitions, we further unfurl various syntactic phenomena that are associated with focus-sensitivity, which are marshaled into 13 descriptive generalizations. These generalizations show focus-sensitivity is a real syntactic phenomenon and cannot be reduced to semantics and phonology. Furthermore, when we treat these descriptive generalizations as diagnostics of focus-sensitivity, it is shown at least the majority of sentence adverbs should be regarded as focus-sensitive. Ignoring these

crucial properties will simply make catastrophically wrong predictions about the syntactic distributions of sentence adverbs cross-linguistically. Then, in the latest leg of our journey (chapter 4), we attempt to go beyond just unraveling the strangeness of the syntax of sentence adverbs by formal theorizing. It is shown the task is not an impossible one, in light of the Agree theory, which implicitly allows the option of delayed syntactic operations. It was shown the theory has the basic proper tools for accounting for various strange properties of sentence adverbs, especially the properties of focus-sensitivity, when focusing adverbs and sentence adverbs alike are treated as 'inflectional affixes writ large'. Many properties that govern the syntactic distributions of focusing adverbs and sentence adverbs are shown to be governed by the same or similar principles to those that govern well-known operations Agree, Pied-pipe, and Merge. As a result, sentence adverbs do not really seem so strange from our perspective, and we are in a position to say that a large part of their apparently puzzling properties are exactly the core properties associated with inflectional morphology and A′-dependency.

5.2 Theoretical consequences

Our application of the Agree theory to the syntax of sentence adverbs has some general consequences for the architecture of grammar.

5.2.1 Support for the Agree theory (as opposed to the Checking theory)

As the title of this thesis implies, our major concern is to show that the syntax of sentence adverbs is but one of the manifold linguistic phenomena that should best be dealt with by the Agree theory, to the extent that no competing theories are shown to be better-equipped at capturing various syntactic dependency relationships. In chapter 4 we do have adopted the Agree theory extensively to deal with the syntactic distributions of sentence adverbs. Consequently, if our approach is on the right track, it should provide further support for the Agree theory. Indeed, we have shown that (i) there is no need to resort to feature movement and spec-head configurations to account for the syntax of FAs and SAs; (ii) even the effects of pair-Merge and the operation SIMPL, at least in the cases we have discussed, can be totally derived from fundamental components of the Agree theory.

5.2.2 The NS-Σ mapping is straightforward (there is no syntax-semantics mismatch)

In chapter 1, we noted that the major theories of adverbial syntax on the market treat sentence adverbs as TP or *v*P-level adjuncts or specifiers, without considering the fact that such

analyses lead to syntax-semantics mismatch problems. These analyses entail either (i) that syntax-semantics mapping is quite free, so a given adverb can attach to any phrase as long as some loose semantic constraints are satisfied (in semantics-oriented approaches), or (ii) that semantically-vacuous syntactic operations, such as head movements, are ubiquitous (in syntax-oriented approaches). The common theoretical problem of these approaches is that the syntax is too free. Either the adverbs can be freely generated by syntax, or that verbal heads can freely move to higher functional heads. There are no well-motivated independent reasons for these freedoms. Furthermore, they also fail to properly describe all the relevant important empirical facts. As we have shown in chapter 2 and 3, sentence adverbs have properties of C^0 expressions, yet their syntax is like those of focus-sensitive adverbs. Simply treating sentence adverbs as TP or vP-level adjuncts or specifiers runs afoul of these facts. Our Agree approach is able to overcome these problems in one fell swoop. In our approach, the NS-Σ mapping is quite straightforward. Sentence adverbs are the 'reflexes' of covert C operators, which appear in non-C positions because they are derived by Agree operations between a C head and a lower functional or lexical head. There are no more semantically vacuous head movements, nor free syntax-semantics mapping. We are able to retain a restrictive syntactic theory in which specific, well-motivated syntactic principles determine syntactic derivations. The apparent freedom of syntactic distributions of FAs and SAs is resulted from the indeterminacy of the focus component a focus-sensitive operator will choose in a given sentence, a highly constrained, albeit sometimes somewhat idiosyncratic distributions of uninterpretable features in a given sentence with a given information structure, and the possibility of adverb-topicalization.

5.2.3 The purpose of Agree is to accommodate the duality of semantics

Another consequence our approach is that it clearly supports Chomsky's (2004 *et seq.*) view that the purpose of uninterpretable features and Agree is to satisfy the duality of semantics. More specifically, they serve as mechanisms of displacement (internal Merge), and the purpose of the latter is to yield discourse-related properties and scopal effects.[1] External Merge, on the other hand, is required for the purpose of realizing θ-relations, which presumably involve a different kind of feature operation.[2] Our analyses supports this view in that they unequivocally show that FAs and SAs also have this duality property, since (i) the *probe* bears an interpretable operator-related feature, (ii) generally, the *goal* bears a θ-related feature, as well as a focus-related feature. It is only with the presence of these features that Agree and concomitant feature operations Pied-pipe, delayed-Merge, and internal Merge apply and yield proper legible

---

[1] See also Miyagawa (2010) for relevant discussion.
[2] See Pesetsky and Torrego (2006) for discussion.

configurations for the Σ component. In other words, if there is no duality of semantics involved, we wouldn't find FAs and SAs to have the focus-sensitive syntactic properties we have discussed extensively in chapter 3.[3]

5.2.4 An updated trinity of syntax: external Merge, internal Merge, and delayed-Merge

On the syntax side, our approach shows not only internal Merge is employed for the purpose of non-θ-related aspects of semantics, delayed-Merge (and the morphosyntactic process Inflect) is also a key syntactic ingredient. On the other hand, the operation pair-Merge has been shown not adequate for the purpose of accounting for the syntax and semantics of FAs and SAs. Consequently, we have an updated trinity of Merge operations: external Merge, internal Merge, and delayed-Merge, the last one replacing pair-Merge. This innovation about how Merge works allows simple accounts for assorted 'late insertion effects' and cases of apparent syntax-semantics mismatches, which have generally been sidestepped in syntactic theories that do not allow late insertion.

One such case we have discussed in chapter 2 and 4 is right-adjunction of various adverbs to verbal (and sometimes adjectival) heads in languages such as English and French. Although there have been forceful arguments against verb movement analyses, such as those found in Iatridou (1990) and Williams (1994, 2000), the alternative right-adjunction analysis they presented have some serious problems at that time. The reason is that the technology at that time only allows $[_V$ V Adv$]$ sequence to be construed as complex predicates that involve morphological incorporation, due to inherent constraints imposed by the Checking theory and the questionable structure-preserving principle that stipulates only YP can adjoin to XP and only Y^0 can adjoin to X^0 in overt syntax. Pollock (1997: 246ff) rightly shows that these predictions are not borne out. There are indeed cases where adverbial phrases composed of two or more words following the verb, and cases where the adverb cannot be considered as part of a complex predicate for semantic reasons. However, our current approach easily overcomes these problems while still adopting the right-adjunction approach. In our Agree theory, a $[_V$ V AP$]$ sequence do not need to involve complex predication, since delayed-Merge is a legitimate option, and adjoining an XP to Y^0 also doesn't violate any principles in the Agree framework.

Other word-level cases of apparent syntax-semantics mismatch that seem to be readily resolved in a similar fashion are as follows:

---

[3] A consequence of this consequence is that uninterpretable features cannot be the source of different parametric settings in languages, since the former are required for semantic purposes. Presumably semantics, as well as 'the language of thought', is uniform cross-linguistically (Chomsky 2005 *et seq*).

(30) a. Frequency adverbs that are realized as adjectives attached to nouns. (e.g. *An occasional sailor strolled by*).[4]

b. Intensives that attach to nouns, VP, or particles (*I have to maw the <u>fucking</u> lawn. He <u>fucking</u> ate the whole goddamn thing. Get her <u>the hell</u> out of there!*).[5]

In these examples, the underlined expressions are adjuncts syntactically, yet their semantic scope is wider than the constituent they attach to. Our Agree approach readily provides a natural syntactic account for these expressions, though the specifications of the loci of uninterpretable features could be (slightly) different in these cases, a detail that awaits further investigations.[6]

A number of other cases that have been labeled as 'late insertion effects' of adjuncts will also have to be accounted for by our delayed-Merge approach, since pair-Merge is eliminated, such as the following examples (Lebeaux 1988, Chomsky 1995: 204, 2004):

(31) a. *Which claim [that John$_i$ was asleep] was he$_i$ willing to discuss?

b. Which claim [that John$_i$ made] was he$_i$ willing to discuss?

The presence of Condition C effect in (31a) and the absence of it in (31b) have been analyzed as adjuncts being merged counter-cyclically. Under our approach, the adjunct *that John made* undergoes delayed-Merge because it is the reflex of Agree that applies after the subject of the main clause (*he*) was merged, and the adjunct is presumably associated with a discourse-related functional head expression topic-related semantics. This seems plausible, but I will leave the details open.[7]

5.2.5 *Narrow syntax* is not so narrow: support for a fine-grained, sub-modular view of NS

Our Agree plus delayed-Merge analyses of FAs and SAs also have consequences on the placement of morphosyntactic operations in the architecture of grammar. Under our analyses, the

---

[4] Previously discussed by Bolinger (1967), Stump (1981), Larson (1999), Zimmerman (2003), Potts (2005), and Morzycki (2008).

[5] Recent discussions of semantic and syntactic properties *fucking* can be found in Potts (2005), Beaver and Clark (2008: 74), Morzycki (2008). Discussions of syntactic properties of *the hell* in example (30b) can be found in Hoeksema and Napoli (2008).

[6] These expressions are perhaps all focus-sensitive. See Beaver and Clark (ibid) on *fucking* as an FSE.

[7] A number of other cases of apparent 'late insertions' have been discussed in the literature, including extrapositions (Fox and Nissenbaum 1999) and comparative constructions (Bhatt and Pancheva 2004). According to F&N and B&P, these constructions involve covert movement (by QR) of an expression followed by overt adjunction/merger to the covertly moved material at its landing site. However, as Chomsky (2004) convincingly points out, there are conceptual problems with these approaches, and there are alternative approaches available without such complexations.

Agree operation is not a cut-off point between narrow syntax and morphology, since delayed-Merge is clearly a syntactic operation that may involve phrasal linguistic units. This implies that Inflect should also be treated as an NS operation, instead of a purely morphology or phonology (Φ) component[8], due to the fact that both delayed-Merge and Inflect serve to realize syntactic feature valuation. What distinguishes Inflect from delayed-Merge is the size of linguistic units that are involved. The former involves combinations of sub-word-level linguistic units, while the latter involve combinations of word-level and phrase-level linguistic units. This partially unified view of phrase-level syntax and word-level syntax runs afoul of the Lexical Integrity Hypothesis/Atomicity Thesis (see §4.1.2), and is also contra to Chomsky's (2001: 5) view that Inflect is an operation in 'phonology', but is very close to the view of the architecture of grammar recently developed by Ackema and Neeleman (2004, 2007), according to which 'Phrasal Syntax' and 'Word Syntax' are submodules of the syntactic macro-module. The operations in the syntactic macro-module all involve combinations of linguistic units that are triggered by syntactic feature operations, but they can differ in whether the linguistic units involved are phrase-level, word-level or sub-word-level expressions, and at different levels different idiosyncratic grammatical principles may have their say.

I will leave open the details of this 'submodular' view of the NS component, which is a vast terrain, but would like to point out some relevant facts that support this view. In a theory of grammar that treats delayed-Merge as an NS operation, and Inflect as a morphology/phonology (Φ) operation, one expects that totally different formal features are involved in these two operations, since they belong to different components. However, this is not true. A striking example is various means languages use to express exclusive identification. In English, the relevant feature is realized by the delayed-Merge of the focusing adverb *only*. However, in other languages exactly the same feature is realized by Inflect: witness the verbal suffix *-dak* in Cantonese, and nominal suffixes *-dake* and *-man* in Japanese and Korean, respectively. Here we see that the same formal feature is involved, but both delayed-Merge and Inflect operations are employed. Assuming a given formal feature involves operations in a given component, we should regard both delayed-Merge and Inflect as operations of the NS component.[9,10]

---

[8] Operations in this component have been previously regarded as 'post-syntactic' (see Embick and Noyer 2001 and references cited there). In Chomsky's current version of the architecture, which Agree theory is based upon, there is no such thing as post-syntactic operations, since NS, Φ, and Σ proceed in parallel.

[9] This also conforms to Chomsky's (2001 *et seq.*) view that NS is basically uniform cross-linguistically, according to which we wouldn't expect that the same feature operation is found in the NS in one language but in Φ in another language.

[10] Note our view (and Ackema and Neeleman's) is different from the view that there are basically no differences between word-syntax and phrase-syntax. For an analysis of focusing affixes in Japanese and Korean under such a view, based on the Checking theory and the Antisymmetry theory, see Koopman (2005). In addition to not being compatible with the Agree theory, it is not clear how such an analysis can capture all the well-established differences between word-level syntax and phrase-level syntax.

5.3 Outlook

The empirical data and the theory provided in this work have provided a general picture of the syntax of sentence adverbs. There are, however, many details and related issues of general interest that regrettably have to be left to the side. In what follows I will summarize what those issues are.

5.3.1 Semantics: mood and focus

One major descriptive set of data featured in this thesis is the focus-sensitivity property of sentence adverbs. However, except Beaver and Clark (2008), so far there have been no real efforts of semantic analyses of this property of sentence adverbs, let alone the connections between semantics and syntax. Our discussion in chapter 3 and 4 has only scratched the surface, since it is based on largely intuitive definitions of focus and focus-sensitivity. In order to correctly characterize the syntax of focus-sensitivity, we need to have a better understanding of the semantics of focus-sensitivity and its relation to the semantics of mood elements. It also goes without saying that the semantics of mood and mood-related expressions themselves is also essential for a proper syntactic theory of sentence adverbs.

5.3.2 Syntax: the nature of locality in Agree operations and relevant issues

Chapter 4 witnesses a number of locality effects that are either relativized to specific lexical items or languages. In addition, the fact that Agree (including Match and Value) between two features at expressions located in different phases show that a mechanism is required to allow Agree to cross a phase boundary. These all show that the theory still leaves many details to be worked out with respect to locality. In addition, we have only touched on some issues relevant for determining locality, such as pied-piping and multiple Agree. These are all very general issues, which involve NS but also seem to involve CI and SM interface factors. A more refined study of locality effects and refinements of theories of locality are thus unavoidable in light of these facts.

5.3.3 The Φ component: prosody, weight, and parenthetical expressions

There are a number of factors involved in the syntax of FAs and SAs which should be regarded as belong to the phonological component (Φ). These include the relation between NS and sentential prosody (which is not always a good diagnostic for semantic focus), the syntax of

parenthetical expressions marked by idiosyncratic syntactic syntax and prosody (see Haegeman (1991) and Arnold (2007)), the fact that phonological "weight" plays a role in determining the surface position of linguistic expressions (see Ernst 2002 for some discussions), and the fact that focus on the auxiliary verb plays a role in adverbial syntax (Baker 1971, 1981, Ernst 1983, Abels 2003). These factors either rely on syntactic features as input, or have a life of their own, and are still open to investigation.

5.3.4 Syntax beyond Agree: mono-clausal vs. bi-clausal structures

As has been discussed in §2.1, there are a number of constructions that resemble the bulk of data discussed in this thesis which I cannot go into details. These are bi-clausal sentences with mood elements serving as predicate expressions, which include the following examples:

(32) a. Predicative adjectives selecting clausal complements (*Mary is likely to win, He was reluctant to answer the question, It's possible that John saw Mary, It's weird that he knows so much, I feel confident leaving the car with you*).
  b. Modal nouns selecting clausal complements (*There is a possibility that Mary will smile*).

These bi-clausal sentences with predicative mood expressions resemble mono-clausal sentences with sentence adverbs in that semantically the predicative elements and adverbs all express either the speaker's attitude of a proposition or the mental state. Syntactically, however, these two types of sentences are clearly distinct, since, as have seen in chapter 2, sentence adverbs display properties of C^0 elements, but the expressions in (32) do not. Unlike the former, the latter can always be negated or questioned, and can sometimes license NPI elements:

(33) a. Mary is not likely to win.
  b. He was not reluctant to answer the question.
  c. It's not possible that John saw Mary.
  d. There is no possibility that Mary will smile.
(34) a. Is Mary likely to win?
  b. Was he reluctant to answer the question?
  c. Is it possible that John say Mary?
  d. Is there a possibility that Mary will smile?
(35) a. I am surprised that he ever speaks to her.
  b. It is surprising that he ever speaks to her.

The semantic similarity and syntactic disparity between the two types of constructions point us to some possibility lines of research. For example, should we assume that sentences with the same semantic structure are mapped to the same syntactic structure? If not, how much can two sentences with the same semantic structure differ syntactically from each other? Do mono-clausal sentences with sentence adverbs and bi-clausal sentences with predicative mood expressions have the same semantic structure?

5.3.5 Syntax beyond Agree: adverbs vs. modal/mood auxiliaries, verbs, and particles

There are several types of mono-clausal sentences that resemble sentences with SAs semantically but are apparently distinct syntactically. Those are sentences with modal auxiliaries, modal verbs, and mood particles. It is clear that understanding the syntax and semantics of these sentences bears on better understanding of sentence adverbs. In this section I will sketch some basic properties that distinguish them syntactically from SAs.

Modal auxiliaries have the following distinctive syntactic properties:

(36) *Syntactic properties of modal auxiliaries*
　　a. At least in languages like English, in declarative sentences they always occur between the subject and the main verb. (*John might know you. *Might John know you.*)
　　b. Their semantic scope vis-à-vis negation is determined by idiosyncratic factors. (*He cannot go. He should not go.*)[11]
　　c. They can undergo T-to-C movement. (*Could he be right? Might he know it?*)
　　d. They can be 'intensified' or 'toned down' by specific adverbs. (*He couldn't possibly have done it by himself, He may very well know Peter. This can all too easily become an addition. The meeting must surely be over by now. The meeting may possibly be over by now. Could you possibly come a little earlier next week?*)[12]

Modal verbs have some properties that overlap with those of modal auxiliaries, but also have some distinctive properties of their own.

(37) *Syntactic properties of modal verbs*
　　a. In declarative sentences they always occur between the subject and the verb.
　　　(*It appears John likes Mary. *Appears it John likes Mary.*)
　　b. Tense/aspectual marking is possible. (*It seemed like a good idea at that time. Bill might*

---

[11] See Iatridou and Sichel (2009) for discussion.
[12] See Anand and Brasoveanu (2010) and references cited there for discussion.

have seemed like a thief. There had to be at least a hundred people there.[13])

c. They can follow modal/aspectual auxiliaries, with the latter taking wide scope. (*It would seem Mary was right. Bill might have seemed like a thief.*)

d. Alternatively, the lower modal verbs take wide scope. (*Bill can't ever seem to find a good jacket*. French: *Il a pu pleuvoir. (it has might-*PPART *rain)* 'It might have rained.' *Il a dû manger. (he has must-*PPART *eat)* 'He must have eaten.')[14]

e. They can take clausal complements. (See Butler (2006))

Mood particles, such as those found in Sino-Tibetan languages, have the following syntactic properties:

(38) *Syntactic properties of mood particles*

a. They occur in the sentence-final position in an SVO language.

b. They can be targets of φ-feature agreement in some languages (see Mei 1996 and references cited there).

c. They exhibit selection restrictions with sentence adverbs (see §2.3.2.2 and §2.4).

Clearly, these expressions are closely related both syntactically and semantically to sentence adverbs. They either have similar semantic functions or interact with sentence adverbs that occur in the same clause. Only by understanding the syntax and semantics of all of these expressions coherently can be construct a proper theory for the cartography project/left-periphery and more specifically, the place of sentence adverbs in the architecture of grammar.

5.3.6 Intra-linguistic and cross-linguistic variations

Throughout this thesis we have seen that although sentence adverbs can be defined by clear-cut syntactic criteria, they are still a heterogeneous group in terms of intra-linguistic and cross-linguistic variations. Sentence adverbs may thus be an ideal testing ground for theories of linguistic variations. For example, we have suggested above that the NS component may be much richer than previously thought, since delayed-Merge and Inflect arguably both happens at NS. Does this mean that NS itself is subject to parametric variation (contra to current minimalist views)? Or that it is the mapping from NS to the Φ component that is the locus of parametric

---

[13] The tense of the modal actually comes from the proposition that the modal modifies. See Stowell (2004) for discussion.

[14] See Langendoen (1970) on the 'can't seem to' construction. See Butler (2006) for some general discussion of modal verbs.

variation? The operation *feature valuation* that triggered delayed-Merge and Inflect also sees lots of intra-linguistic and cross-linguistic variations. Although we have seen that languages are to some extent uniform in that semantic focus plays a role in determining the *goal* of the Agree operation, many micro-variations exist. What exactly are the factors involved in these variations? Can we predict what intra-linguistic and cross-linguistic variations are possible and what are not? These are perhaps the most difficult and most important questions of a syntactic theory, without understanding which we cannot say we have a proper theory of sentence adverbs.

5.3.7 The Kingdom of Agree

What have been unraveled in this study may serve as a useful stepping-stone toward further studies of other domains of language, especially the functional categories. The Agree theory has already proven useful to capture a large number of syntactic phenomena and make the very simple generalizations. Refining the theory and applying it to account for the syntax of sentence adverbs allows us to capture even more empirical facts with simple generalizations, and with theoretical coherence. Based on our discoveries, there could all too easily be many other hidden syntactic dependency relations that are waiting to be discovered. If the Agree theory is a useful tool for discovering these dependency relations, there is no reason not to use it. Now, we have found that focusing adverbs, sentence adverbs, and inflectional morphology can be accounted for in an elegant fashion in the theory. It remains to see whether other classes of adverbial adjuncts and derivational morphology have any bearing on the Agree theory, and how other cases of syntax-semantic mismatch (e.g. modal verbs and modal auxiliaries) are to be accounted for by the Agree theory or some very different theory.

References

Abels, Klaus. 2003. Auxiliary adverb word order revisited. In *Proceedings of UPenn Colloquium on Linguistics* 26: 1-15.

Abney, Steven. 1987. The English noun phrase in its sentential aspect. PhD dissertation. MIT.

Aboh, Enoch O. 2010. Information structuring begins with the numeration. *Iberia* 2: 12-42.

Ackema, Peter and Ad Neeleman. 2004. *Beyond morphology: interface conditions on word formation*. Oxford: Oxford University Press.

Ackema, Peter and Ad Neeleman. 2007. Morphology ≠ syntax. *The Oxford handbook of linguistic interfaces*. Oxford: Oxford University Press.

Adger, David and Josep Quer. 2001. The syntax and semantics of unselected embedded questions. *Language* 77: 107-133.

Adger, David. 2003. *Core syntax: a Minimalist approach.* Oxford: Oxford University Press.

Åfarli, Tor A. 1997. Dimensions of phrase structure: the representation of sentence adverbials. *Motskrift* 2.97: 91-113, Norwegian University of Science and Technology (NTNU), Trondheim.

Aldridge, Edith. 2010. Clause-internal *wh*-movement in Archaic Chinese. *Journal of East Asian Linguistics* 19: 1-36.

Alexiadou, Artemis. 1997. *Adverb placement: a case study in antisymmetric syntax*. Amsterdam: John Benjamins.

Alexiadou, Artemis. 2002. The syntax of adverbs: puzzles and results. *Glot International* 6 2/3: 33-54.

Alexiadou, Artemis and Elena Anagnostopoulou. 1998. Parametrizing AGR: word order, V-movement, and EPP-checking. *Natural Language and Linguistic Theory* 16: 491-539.

An, Young-ran. 2007. Re extrinsic *–tul*. Ms. Stony Brook University.

Anand, Pranav and Adrian Brasoveanu. 2010. Modal concord as modal modification. In *Proceedings of Sinn und Bedeutung* 14: 19-36.

Anderson, Stephen R. 1972. How to get *even*. *Language* 48: 893-906.

Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.

Andrews, Avery. 1983. A note on the constituent structure of modifiers. *Linguistic Inquiry* 14: 695-697.

Aoun, Joseph and Y.-H. Audrey Li. 1989. Constituency and scope. *Linguistic Inquiry* 20: 141-172.

Aoun, Joseph and Y.-H. Audrey Li. 1993a. *Syntax of scope*. Cambridge, Mass.: MIT Press.

Aoun, Joseph and Y.-H. Audrey Li. 1993b. *Wh*-elements in-situ: syntax or LF? *Linguistic Inquiry* 24: 199-238.

Aoyagi, Hiroshi. 1998. On the nature of particles in Japanese and its theoretical implications. PhD dissertation. USC.

Arnold, Doug. 2007. Non-restrictive relatives are not orphans. *Journal of Linguistics* 43: 271-309.

Aronoff, Mark. 1994. *Morphology by itself: stems and inflectional classes*. Cambridge, Mass.: MIT Press.

Asher, Nicholas. 1993. Reference to abstract objects in discourse. Dordrecht: Kluwer Academic Publishers.

Baker, Carl L. 1971. Stress level and auxiliary behavior in English. *Linguistic Inquiry* 2: 167-181.

Baker, Carl L. 1981. Auxiliary-adverb word order. *Linguistic Inquiry* 12: 309-315.

Baker, Mark. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

Bayer, Josef. 1996. *Directionality and logical form: on the scope of focusing particles and wh-in-situ*. Dordrecht: Kluwer Academic Publishers.

Bayer, Josef. 1999. Bound focus or how association with focus be achieved without going semantically astray? In *The grammar of focus*, ed. by Georges Rebuschi and Laurice Tuller, 55-82. Amsterdam: John Benjamins.

Beaver, I. David and Brady Z. Clark. 2008. *Sense and sensitivity: how focus determines meaning*. Chichester: Wiley-Blackwell.

Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14: 1-56.

Beck, Sigrid and Shin-sook Kim. 1997. On *wh-* and operator scope in Korean. *Journal of East Asian Linguistics* 6: 339-384.

Bellert, Irena. 1977. On semantic and distributional properties of sentenctial adverbs. *Linguistic Inquiry*. 8: 337-351.

Belletti, Adriana. 1990. *Generalized verb movement*. Turin: Rosenberg & Sellier.

Belletti, Adriana. 1994. Verb positions: evidence from Italian. In *Verb movement*, ed. by David Lightfoot and Nobert Hornstein, 19-40. Cambridge: Cambridge University Press.

Belletti, Adriana. 2004. Aspects of the low IP area. In *The structure of CP and IP*, ed. by L. Rizzi, 16-51. Oxford: Oxford University Press.

Belletti, Adriana, ed. 2004. *The cartography of syntactic structures.* Vol. 3. *Structures and beyond*. Oxford: Oxford University Press.

Bennett, Jonathan. 1988. *Event and their names*. Indianapolis: Hackett Publishing Company.

Beninca, Paola and Nicola Munaro, ed. To appear. *The cartography of syntactic structures.* Vol. 5. *Mapping the left periphery.* Oxford: Oxford University Press.

Bhatt, Rajesh and Roumyana Pancheva. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35: 1-45.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written English.* Harlow, Essex: Pearson Education Limited.

Biberauer, Theresa, Anders Holmberg, and Ian Roberts. 2009. Linearization and the architecture of grammar: A view from the Final-over-Final constraint. Paper presented at *NELS* 40, MIT, Cambridge, MA.

Bisang, Walter. 1998. Adverbiality: the view from the Far East. In *Adverbial constructions in the languages of Europe*, ed. by Johan van der Auwera, 641-812. Berlin: Mouton de Gruyter.

Bobaljik, Jonathan. 1995. Morphosyntax: the syntax of verbal inflection. PhD dissertation. MIT.

Bobaljik, Jonathan. 1999. Adverbs: the hierarchy paradox. *Glot International* 4 9/10: 27-28.

Bobaljik, Jonathan. 2008. Where is Phi? Agreement as a postsyntactic operation. In *Phi theory: Phi-features across modules and interfaces*, ed. by Daniel Harbour, David Adger, and Susana Béjar, 295-328. Oxford: Oxford University Press.

Bobaljik, Jonathan and Dianne Jonas. 1996. Subject positions and the role of TP. *Linguistic Inquiry* 27: 195-236.

Boeckx, Cedric. 2001. Scope reconstruction and A-movement. *Natural Language and Linguistic Theory* 19: 503-548.

Boeckx, Cedric. 2003. Review of T. Ernst (2002) The syntax of adjuncts, Cambridge University Press. *Studies in the Linguistic Sciences* 32: 95-101.

Boeckx, Cedric. 2008. *Bare syntax.* Oxford: Oxford University Press.

Bolinger, Dwight. 1967. Adjectives in English: attribution and predication. *Lingua* 18: 1-34.

Bonami, Oliver and Danièle Godard. 2008. Lexical semantics and pragmatics of evaluative adverbs. In *Adjectives and adverbs: syntax, semantics, and discourse*, ed. by Louise McNally and Christopher Kennedy, 273-304. New York: Oxford University Press.

Bošković, Željko. 2004. Topicalizaiton, focalization, lexical insertion, and scrambling. *Linguistic Inquiry* 35: 613-638.

Bošković, Željko. 2009. Review of Lisa L.-S. Cheng and Norbert Corver, ed. (2006) Wh-movement: moving on. Cambridge, MA: MIT Press. *Language* 85: 463-468.

Brennan, Jonathan. 2007. *Only* finally. *NYU Working Papers in Linguistics* 1: 1-14.

Bruening, Benjamin. 2001. QR obeys superiority: frozen scope and ACD. *Linguistic Inquiry* 32: 233-273.

Büring, Daniel and Katharina Hartmann. 2001. The syntax and semantics of focus- sensitive

particles in German. *Natural Language & Linguistic Theory* 19: 229-281.

Butler, Johnny. 2003. A minimalist treatment of modality. *Lingua* 113: 967-996.

Butler, Johnny. 2006. The structure of temporality and modality (or, Towards deriving something like a Cinque Hierarchy). *Linguistic Variation Yearbook* 6: 161-201.

Cardinaletti, Anna and Michal Starke. 1999. The typology of structural deficiency: a case study of the three classes of pronouns. In *Clitics in the languages of Europe*, ed. by Henk van Riemsdijk, 145-233. Berlin: Mouton de Gruyter.

Chang, Henry Y. 2006. The guest playing host: adverbial modifiers as matrix verbs in Kavalan. In *Clause structure and adjuncts in Austronesian languages*, ed. by Hans-Martin Gaertner et al., 43-82. Berlin: Mouton de Gruyter.

Chang, Melody Y.-Y. 2001. Functional categories and adverbial expressions: a case study of Paiwan and Tsou. *Report for the National Science Council*.

Chao, Yuen-ren. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.

Cheng, Lisa L.-S. 1991. On the typology of wh-questions. PhD dissertation. MIT.

Cheng, Lisa L.-S. and Rint Sybesma. 2003. Forked modality. *Linguistics in the Netherlands 2003*, 13-23.

Cheung, Lawrence Y.-L. 2009. Dislocation focus construction in Chinese. *Journal of East Asian Linguistics* 18: 197-232.

Choe, Hyon Sook. 1987. An SVO analysis of VSO languages and parameterization: a study of Berber. In *Studies in Berber syntax. Lexical Project Working Paper 14*, ed. by Mohamed Guerssel and Ken Hale, 121-158. Distributed by MITWPL.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, ed. by Roderick Jacobs and Peter S. Rosenbaum, 184-221. Waltham, Mass.: Ginn & Company.

Chomsky, Noam. 1971. Deep structure, surface structure, and semantic interpretation. In *Semantics: an interdisciplinary reader*, ed. by Danny D. Steinberg & Leon A. Jakobovits, 183-216. Cambridge: Cambridge University Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986a. *Knowledge of language: its nature, origin, and use*. New York: Praeger.

Chomsky, Noam. 1986b. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: a life in language*, ed. by Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structure and beyond: the cartography of syntactic structures*, vol 3, ed. by Adriana Belletti, 104-131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1-22.

Chomsky, Noam. 2007. *Approaching UG from below.* In *Interfaces + recursion = language?: Chomsky's Minimalism and the view from syntax-semantics*, ed. by Uli Sauerland and Gans-Martin Gärtner, 1-29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133-166. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2009. Opening remarks. In *Of mind & language: a dialogue with Noam Chomsky in the Basque Country*, ed. by Massimo Piattelli-Palmarini, Juan Uriagereka, and Pello Salaburu, 13-43. Oxford: Oxford University Press.

Chung, Daeho. 2004. Semantics and syntax of correlative adverbs. *Studies in Generative Grammar*. 14.3: 15-28.

Chung, Sandra. 1990. VP's and verb movement in Chamorro. *Natural Language and Linguistic Theory* 8: 559-619.

Chung, Sandra. 1998. *The design of agreement: evidence from Chamorro*. Chicago: University of Chicago Press.

Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239-297.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.

Cinque, Guglielmo, ed. 2002. *The cartography of syntactic structures. Vol*. 1, *Functional structure in DP and IP*. Oxford: Oxford University Press.

Cinque, Guglielmo, ed. 2006. *The cartography of syntactic structures. Vol*. 4, *Restructuring and functional heads*. Oxford: Oxford University Press.

Cinque, Guglielmo and Luigi Rizzi. 2008. The cartography of syntactic structures. *CISCL Working Papers on Language and Cognition 2*: 43-59.

Collins, Christopher. 1988. Conjunction adverbs. Ms. MIT.

Collins, Christopher. 1991. *Why* and *how come. MIT Working Papers in Linguistics* 15: 31-45.

Collins, Christopher. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel D. Epstein and Daniel Seely, 42-64. Malden, Mass.: Blackwell Publishing Ltd.

Corbett, Greville G. 1998. Morphology and agreement. In *The handbook of morphology*, ed. by Andrew Spencer and Arnold M. Zwicky, 191-205. Oxford: Blackwell Publishers Ltd.

Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.

Costa, João. 1996. Adverb positioning and V-movement in English: some more evidence. *Studia Linguistica* 50: 22-34.

Creswell, Cassandre. 2000. The discourse function of verum focus in *wh*-questions. *NELS* 30: 165-179.

Culicover, Peter. 1991. Polarity, inversion, and focus in English. *ESCOL* 8: 46-68.

Den Dikken, Marcel. 2006. Either-float and the syntax of co-or-dination. *Natural Language and Linguistic Theory* 24: 689-749.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.

Di Sciullo, Anna-Maria and Edwin Williams. 1987. *On the definition of word*. Cambridge, Mass.: MIT Press.

Drubig, Hans Bernhard. 1994. Island constraints and the syntactic nature of focus and association with focus. In *Arbeitspapiere des Sonderforschungsbereichs 340: Sprachtheoretische Grundlagen der Computerlinguistik* (Vol. 51). Tübingen/Universität Stuggart.

Egedi, Barbara. 2009. Adverbial (dis)ambiguities. Syntactic and prosodic features of ambiguous predicational adverbs. In *Adverbs and adverbial adjuncts at the interfaces*, ed. by Katalin É Kiss, 103-132. Berlin: Mouton de Gruyter.

É Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74: 245-273.

Embick, David and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32: 555-595.

Emonds, Joseph. 1985. *A unified theory of syntactic categories*. Dordrecht: Foris.

Engdahl, Elisabet, Maia Andréasson and Kersti Börjars 2004. Word order in the Swedish midfield – an OT approach. In *Proceedings of the 20th Scandinavian Conference of Linguistics*, ed. by Fred Karlsson, 1-13. University of Helsinki.

Engels, Eva. 2005. Adverb placement: an Optimality Theoretic approach. PhD dissertation, University of Potsdam.

Ernst, Thomas. 1984. Towards an integrated theory of adverb position in English. PhD dissertation. Indiana University.

Ernst, Thomas. 1983. More on adverbs and stressed auxiliaries. *Linguistic Inquiry* 14: 542-549.

Ernst, Thomas. 1991. On the Scope Principle. *Linguistic Inquiry* 22: 750-756.

Ernst, Thomas. 2002. *The syntax of adjuncts*. Cambridge: Cambridge University Press.

Ernst, Thomas. 2009. Speaker-oriented adverbs. *Natural Language and Linguistic Theory* 27: 497-544.

Ernst, Thomas and Chengchi Wang. 1995. Object preposing in Mandarin Chinese. *Journal of*

East Asian Linguistics 4: 235-260.

Fanselow, Gisbert. 2006. Partial *wh*-movement. In *The Blackwell companion to syntax, vol. 3*, ed. by Martin Everaert and Henk van Riemsdijk, 437-492. Malden, Mass.: Blackwell Publishing Ltd.

Fanselow, Gisbert and Damir Ćavar. 2000. Remarks on the economy of pronunciation. In *Competition in syntax*, ed. by Gereon Müller and Wolfgang Sternefeld, 107-150. Amsterdam: John Benjamins.

Fiengo, Robert, C.-T. James Huang, Howard Lasnik, and Tanya Reinhart. 1988. The syntax of wh-in-situ. *WCCFL* 7: 81-98.

Fox, Danny and Jon Nissenbaum. 1999. Extraposition and scope: a case for overt QR. *WCCFL* 18: 132-144.

Frey, Werner and Karin Pittner. 1998. Zur Positionierung der Adverbiale im deutschen Mittelfeld. *Linguistische Berichte* 176: 489-354.

Frey, Werner. 2010. Ā-Movement and conventional implicatures: about the grammatical encoding of emphasis in German. *Lingua* 120: 1416-1435.

Fukui, Naoki and Yuji Takano. 1998. Symmetry in syntax: Merge and Demerge. *Journal of East Asian Linguistics* 7: 27-86.

Futagi, Yoko. 2004. Japanese focus particles at the syntax-semantics interface. PhD dissertation. Rutgers University.

Grewendorf, Günther. 2005. The discourse configurationality of scrambling. In *The free word order phenomenon: its syntactic sources and diversity*, ed. by Joachim Sabel and Mamoru Saito, 75-135. Berlin: Mouton de Gruyter.

Grimshaw, Jane. 1991. Extended projection. Ms., Rutgers University.

Grimshaw, Jane. 2000. Locality and extended projection. In *Lexical specification and insertion*, ed. by Peter Coopmans, Martin Everaert, and Jane Gramshaw, 115-133. Amsterdam: John Benjamins.

Groat, Erich and John O′Neil. 1996. Spell-Out at the LF-interface. In *Minimal ideas: syntactic studies in the Minimalist framework*, ed. by Werner Abraham, Samuel Epstein, Höskuldur Thráinsson, and C. Jan-Wouter Zwart, 113-139. Amsterdam: John Benjamins.

Guerzoni, Elena. 2003. *Why even ask?* On the pragmatics of questions and the semantics of answers. PhD dissertation. MIT.

Gutzmann, Daniel. 2009. Modal particles, stress, and sentence mood: a use-conditional approach. Paper presented at *40 Jahre Partikelforschung* [40 years of particle research]: 1969–2009. University of Bern.

Haegeman, Liliane. 1991. Parenthetical adverbials: the radical orphanage approach. In *Aspects of modern English linguistics: paper presented to Masatomo Ukaji on his 60th birthday*, ed. by

Shuki Chiba, Akira Ogawa, Yasuki Fuiwara, Norio Yamada, Osamu Koma, and Takao Yagi, 232-254. Tokyo: Kaitakushi.

Haegeman, Liliane. 2000a. Inversion, non-adjacent inversion, and adjuncts in CP. *Transactions of the Philological Society* 98.1: 121- 160.

Haegeman, Liliane. 2000b. Negative preposing, negative inversion, and the Split CP. In *Negation and Polarity*, ed. by Lawrence R. Horn and Yasuhiko Kato, 21-61. New York: Oxford University Press.

Haegeman, Liliane. 2003. Conditional clauses: external and internal syntax. *Mind and Language* 18: 317-339.

Haegeman, Liliane. 2006. Conditionals, factives and the left periphery. *Lingua* 116: 1651-1669.

Haegeman, Liliane. 2010. The internal syntax of adverbial clauses. *Lingua* 120: 628-648.

Haegeman, Liliane and Terje Lohndal. 2010. Negative concord and (multiple) Agree: a case study of West Flemish. *Linguistic Inquiry* 41: 181-211.

Harris, James W. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22: 27-62.

Hendriks, Petra. 2002. "Either" as a focus particle. Ms., University of Gronigen.

Hendriks, Petra. 2004. *Either*, *both*, and *neither* in coordinate structures. In *The composition of meaning: from lexeme to discourse*, ed. by Alice ter Meulen and Werner Abraham, 115-138. Amsterdam: John Benjamins.

Heny, Frank. 1973. Sentence and predicate modifiers in English. In *Syntax and semantics* 2, ed. by John P. Kimball, 217-245. New York: Seminar Press.

Herburger, Elena. 2000. What counts: focus and quantification. Cambridge, Mass.: MIT Press.

Hinterwimmer, Stefan. 2006. The Interpretation of Universally Quantified DPs and Singular Definites in Adverbially Quantified Sentences. *WCCFL* 25: 195-203.

Hiraiwa, Ken. 2005. Dimensions of symmetry in syntax: agreement and clausal architecture. PhD dissertation. MIT.

Hoeksema, Jack & Frans Zwarts. 1991. Some remarks on focus adverbs. *Journal of semantics* 8: 51-70.

Hoeksema, Jack and Donna-Jo Napoli. 2008. Just for the hell of it: a comparison of two taboo-term constructions. *Journal of Linguistics* 44: 347-378.

Höhle, Tilman N. 1992. Über Verum Fokus im Deutschen. *Linguistische Berichte Sonderheft* 4: 112-141.

Hole, Daniel P. 2004. *Focus and background marking in Mandarin Chinese: system and theory behind* cái, jiù, dōu, and yě. London: RoutledgeCurzon.

Holmberg, Anders. 1993. Two subject positions in IP in Mainland Scandinavian. *Working Papers in Scandinavian Syntax* 52: 29-41.

Holmberg, Anders. 2001. The syntax of yes and no in Finnish. *Studia Linguistica* 55: 140-174.

Holmer, Arthur. 2006. Seediq－Adverbial heads in a Formosan language. In *Clause structure and adjuncts in Austronesian languages*, ed. by Hans-Martin Gaertner et al., 83-123. Berlin: Mouton de Gruyter.

Hornstein, Nobert, and Jairo Nunes. 2008. Adjunction, labeling, and Bare Phrase Structure. *Biolinguistics* 2.1: 57-86.

Horvath, Julia. 2006. Pied-piping. In *The Blackwell companion to syntax, vol. 3*, ed. by Martin Everaert and Henk van Riemsdijk, 569-630. Malden, Mass.: Blackwell Publishing Ltd.

Horvath, Julia. 2007. Separating "focus movement" from focus. In *Phrasal and clausal architecture: syntactic derivation and interpretation. In honor of Joseph E. Emonds*, ed. by Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 108-145. Amsterdam: John Benjamins.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. PhD dissertation. MIT.

Huang, C.-T. James. 2003. The distribution of negative NPs and some typological correlates. In *Functional structure(s), form and interpretation: perspectives from East Asian languages*, ed. by Audrey Y.-H. Li and Andrew Simpson, 262-280. London: Routledge.

Huang, C.-T. James and Masao Ochi. 2004. Syntax of *the hell*: two types of dependencies. *NELS* 34: 279-293.

Huang, He. 1990. Changyong fuci gongxian shi de shuenxu [The order of common adverbs]. In *Zhuiyuji*: 494-524. Beijing: Beijing University Publishing.

Iatridou, Sabine. 1990. About Agr(P). *Linguistic Inquiry*. 21: 551-577.

Iatridou, Sabine and Ivy Sichel. 2009. Negative DPs and scope diminishment: some basic patterns. *NELS* 38: 411-424.

Ifantidou-Trouki, Elly. 1993. Sentential adverbs and relevance. *Lingua* 90: 69-90.

Irurtzun, Aritz. 2007. The grammar of focus at the interfaces. PhD dissertation. University of the Basque Country.

Irwin, Patricia. 2009. Polarity and degree in *so totally* constructions. CUNY Syntax Supper.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

Jackendoff, Ray. 1977. *X̄ syntax: a study of phrase structure*. Cambridge, Mass.: MIT Press.

Jacobs, Joachim. 1983. *Fokus und Skalen: Zur Syntax und Semantik der Gradpartikeln im Deutschen*. Tübingen: Niemeyer.

Jaeger, Florian and Michael Wagner. 2003. Association with focus and linear order in German. Ms., Stanford University.

Jayaseelan, Karattuparambil A. 2001. IP-internal topic and focus phrases. *Studia Linguistica* 55: 39-75.

Jespersen, Otto. 1924. *The philosophy of grammar*. London: Allen & Unwin.

Johannessen, Janne Bondi. 2005. The syntax of correlative adverbs. *Lingua*. 115: 419-443.

Jónsson, Gísli Jónsson. 2002. S-adverbs in Icelandic and the feature theory of adverbs. In *Leeds Working Papers in Linguistics and Phonetics* 9: 73-89.

Julien, Marit. 2000. Syntactic heads and word formation: a study of verbal inflection. PhD dissertation. University of Tromsø.

Kaplan, David. 1999. The meaning of *ouch* and *oops*. Explorations in the theory of meaning as use. 2004 version. Ms. UCLA.

Karttunen, Lauri and Stanley Peters. 1979. Conventional implicature. In *Syntax and semantics 11: Presuppositions*, ed. by Choon-kyu Oh and David A. Dinneer, 1-55. New York: Academic Press.

Katzir, Roni. 2011. Morphosemantic mismatches, structural economy, and licensing. *Linguistic Inquiry* 42: 45-82.

Kawamura, Tomoko. 2007. Some interactions of focus and focus sensitive elements. PhD dissertation. Stony Brook University.

Kayne, Richard. 1984. *Connectedness and binary branching*. Dordrecht: Foris.

Kayne, Richard. 1989. Notes on English agreement. *CIEFL Bulletin* 1: 40-67.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1: 128-191.

Keyser, S. Jay. 1968. Review of Sven Jacobson (1964) Adverbial positions in English (Uppsala dissertation). *Language* 44: 357-374.

Kibrik, Aleksandr E. 1994. Archi. In *Indigenous languages of the Caucasus IV: North East Caucasian languages II: presenting the three Nakh languages and six minor Lezgian languages,* ed. by Rieks Smeets, 297-365. Delmar: Caravan Books.

Kim, Soo-won. 1991. Chain scope and quantification structure. PhD dissertation. Brandeis Univeristy.

Kim, Yookyung. 1994. A non-spurious account of 'spurious' Korean plurals. In *Theoretical issues in Korean linguistics*, ed. by Young-Key Kim-Renaud, 303-323. Stanford, Calif.: CSLI Publications.

Kiss, Katalin É. 1988. Identificational focus versus information focus. *Language* 74: 245-273.

Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61: 95-120.

König, Ekkehard. 1991. *The meaning of focus particles: a comparative perspective*. London: Routledge.

Koopman, Hilda. 2005. Korean (and Japanese) morphology from a syntactic perspective. *Linguistic Inquiry* 36: 601-633.

Krifka, Manfred. 1992. A compositional semantics for multiple focus constructions. In *Informationsstruktur und Grammatik*, ed. by Joachim Jacobs, volume Sonderheft 4 of Linguistische Berichte, 17–53. Opladen: Westdeutscher Verlag.

Krifka, Manfred. 2006. Association with focus phrases. In *The architecture of focus*, ed. by Valéria Molár and Susanne Winkler, 105-136. Berlin: Mouton de Gruyter.

Krifka, Manfred. 2007. Basic notions of information structure. In *Working Papers of theSFB632, Interdisciplinary Studies on Information Structure* (ISIS) 6, ed. by Caroline Féry, Gisbert Fanselow, and Manfred Krifka, 13-56. Potsdam: Universitätsverlag Potsdam.

Ladusaw, William A. 1979. Polarity sensitivity as inherent scope relations. PhD dissertation. University of Texas at Austin.

Ladusaw, William A. 1988. Adverbs, negation, and QR. In *Linguistics in the morning calm* 2: 481-488. Seoul: Hanshin Publishing Co.

Laka, Mugarza. 1990. Negation in syntax: on the nature of functional categories and projections. PhD dissertation, MIT.

Langendoen, D. Terence. 1970. The 'can't seem to' construction. *Linguistic Inquiry* 1: 25-35.

Lasnik, Howard. 1999a. Chains of arguments. In *Working minimalism*, ed. by Samuel D. Epstein and Nobert Hornstein, 189-215. Cambridge, Mass.: MIT Press.

Lasnik, Howard. 1999b. *Minimalist analysis*. Malden, Mass.: Blackwell Publishing Ltd.

Larson, Richard. 2004. Sentence-final adverbs and "scope". *NELS* 34: 23-43.

Larson, Richard. 1999. Semantics of adjectival modification. LOT Winterschool class notes, Amsterdam.

Law, Paul. 2008. The *wh/q*-polarity adverb *daodi* in Mandarin Chinese and the syntax of focus. *The Linguistic Review* 25: 297-345.

Lebeaux, David. 1988. Language acquisition and the form of grammar. PhD. Dissertation. University of Massachusetts. Revised and extended edition, Amsterdam: John Benjamins, 2000.

Lechner, Winfried. 2000. Bivalent coordination in German. *Snippets* 1: 11-12.

Lee, Youngjoo. 2004. The syntax and semantics of focus particles. PhD dissertation. MIT.

Lee, Youngjoo. 2005. Exhaustivity as agreement: the case of Korean *man* 'only'. *Natural Language Semantics* 13: 169-200.

Li, Charles N. and Sandra A. Thompson. 1981. *Mandarin Chinese: a functional reference grammar.* Berkeley: University of California Press.

Li, Jie. 2005. Shi lun xiandai hanyu yuqi fuci zhuangyu de xinxi gongneng [On the informative function of modal adverb as adverbial in Chinese]. *Journal of Xinjiang University (Philosophy, Humanities & Social Sciences)* 33.2: 137-141.

Li, Y.-H. Audrey. 1998. Argument determiner and number phrases. *Linguistic Inquiry* 29:

693-702.

Lieber, Rochelle. 1992. *Deconstructing morphology: word formation in syntactic theory*. Chicago: University of Chicago Press.

Longobardi, Giuseppe. 1991. In defense of the correspondence hypothesis: Island effects and parasitic constructions in logical form. In *Logical structure and linguistic structure. Cross-linguistic perspectives*, ed. by C.-T. James Huang & Robert May, 149-196. Dordecht: Kluwer.

Lu, Jian-ming. 1980. Hanyu kouyu jufa li de yiwei xianxiang [Dislocation in the syntax of colloquial Chinese]. *Zhongguo Yuwen* 1: 28-41.

Lü, Shu-xiang. 1980. *Xiandai hanyu babai ci* [800 words of contemporary Chinese]. Beijing: Shangwu Publishing Co.

Lü, Shu-xiang. 1985. Yiwen, kending, fouding [Question, negation, and assertion]. *Zhongguo Yuwen* 187: 241-250.

Luo, Xiao-ying and Shao Jing-min. 2006. The semantic exploration on the adverb *ke* and its pragmatic explanation. *Journal of Jinan University* (*Philosophy and Social Sciences*) 121: 102-107.

Lyons, John. 1995. *Linguistic semantics: an introduction*. Cambridge: Cambridge University Press.

Marantz, Alec. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st annual Penn Linguistics Colloquium*, ed. by Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams, 201-225. University of Pennsylvania.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37: 69-109.

May, Robert. 1977. The grammar of quantification. PhD dissertation. MIT.

McCawley, James D. 1988. *The syntactic phenomena of English*. Chicago: University of Chicago Press.

McConnell-Ginet, Sally. 1982. Adverbs and logical form. *Language* 58: 144-184.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry*. 37: 69-109.

McClosky, James. 1996. On the scope of verb-movement in Irish. *Natural Language and Linguistic Theory* 14: 46-104.

McClosky, James. 1997. Subjecthood and subject positions. In *Elements of grammar*, ed. by Liliane Haegeman, 197-235. Dordrecht: Kluwer Academic Publishers.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation in the minimalist program*, ed. by Samuel David Epstein and Daniel Seely, 184-226. Oxford: Blackwell Publishers Ltd.

Mei, Kuang. 1996. Dulongyu de juweici yanjiu [A study of sentence-final particles in Trung].

Yuyen Yenjou 30: 151-175.

Meinunger, André. 2006. Interface restrictions on verb second. *The Linguistic Review* 23: 127-160.

Miller, Philip H. 1992. *Clitics and constituents in phrase structure grammar*. New York: Garland.

Miyagawa, Shigeru. 2010. *Why Agree? Why Move? Unifying agreement-based and discourse-configurational languages*. Cambridge, Mass.: MIT Press.

Molnár, Valéria and Susanne Winkler. 2010. Edges and gaps: contrast at the interfaces. *Lingua* 120: 1392-1415.

Morzycki, Marcin. 2008. Nonrestrictive modifiers in non-parenthetical positions. In *Adjectives and adverbs: syntax, semantics, and discourse*, ed. by Louise McNally and Christopher Kennedy, 101-122. Oxford: Oxford University Press.

Neeleman Ad and Hans van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11: 137-189.

Nevins, Andrew. Derivations without the Activity Condition. *MIT Working Papers in Linguistics* 49: 283-306.

Nilsen, Øystein. 2001. Adverb order in type-logical grammar. In *Proceedings of the Amsterdam Colloquium 2001*, ed. by Rob van Rooy and Martin Stokhof, 156–161.

Nilson, Øystein. 2003. Eliminating positions: syntax and semantics of sentence modification. PhD dissertation. Utrecht University.

Nilsen, Øystein. 2004. Domains for adverbs. *Lingua* 114.6: 809-847.

Nishigauchi, Taisuke. 1986. Quantification in syntax. PhD dissertation. University of Massachusetts.

Ouhalla, Jamal. 1990. Sentential negation, Relativized Minimality and the aspectual status of auxiliaries. *The Linguistic Review* 7: 183-231.

Parsons, Terence. 1990. *Events in the semantics of English: a study in subatomic semantics*. Cambridge, Mass.: MIT Press.

Paul, Waltraud. 2009. Constituent disharmony: sentence-final particles in Chinese. To appear in *Cambridge Occasional Papers in Linguistics,* vol 5.

Penka, Doris. 2007. Negative indefinites. PhD dissertation. University of Tübingen.

Pesetsky, David. 1989. Language particular processes and the earliness principle. Ms. MIT.

Pesetsky, David and Esther Torrego. 2001. T-to-C movement: causes and consequences. In *Ken Hale: a life in language*, ed. by Kenstowicz Michael, 355-426. Cambridge, Mass.: MIT Press.

Pesetsky, David and Esther Torrego. 2004. Tense, case, and the nature of syntactic categories. In *The syntax of time*, ed. by Jacqueline Guéron and Jacqueline Lecarme, 495-537. Cambridge,

Mass.: MIT Press.

Pesetsky, David and Esther Torrego. 2006. Probes, goals and syntactic categories. In *Proceedings of the Seventh Conference on Psycholinguistics*, ed. by Yukio Ttsu, 25-60. Tokyo: Hituzi Syobo.

Pesetsky, David and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: Syntactic derivation and interpretation. In honor of Joseph E. Emonds,* ed. by Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262-294. Amsterdam: John Benjamins.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365-424.

Pollock, Jean-Yves. 1997. Notes on clause structure. In *Elements of grammar*, ed. by Liliane Haegeman, 237-279. Dordrecht: Kluwer Academic Publishers.

Postal, Paul, M. 1974. *On raising*. Cambridge, Mass.: MIT Press.

Potts, Christopher. 2005. *The logic of conventional implicatures*. New York: Oxford University Press.

Pires, Acrisio. 2001. The syntax of gerunds and infinitives: subject, case and control. PhD dissertation. University of Maryland, College Park.

Pullum, Geoffrey K. and Rodney Huddleston. 2002. Adjectives and adverbs. In *The Cambridge grammar of the English language*, ed. by Geoffrey K. Pullum and Rodney Huddleston, 525-595. Cambridge: Cambridge University Press.

Quer, Josep. 2009. Twists of mood: The distribution and interpretation of indicative and subjunctive. *Lingua* 119: 1779-1787.

Qi, Chun-hong. 2006. Xiandai hanyu yuqi fuci yanjiu [A study of mood adverbs in modern Chinese]. PhD dissertation. HuaZhong Normal University.

Radford, Andrew. 1988. *Transformational Grammar: a first course*. Cambridge: Cambridge University Press.

Radford, Andrew. 2004. *English syntax: an introduction*. Cambridge: Cambridge University Press.

Reinhart, Tanya. 1976. The syntactic domain of anaphora. PhD dissertation. MIT.

Reinhart, Tanya 1984. *Anaphora and Semantic Interpretation.* Chicago: University of Chicago Press.

Reinhart, Tanya. 1995. *Interface strategies*. OTS Working Papers. Utrecht University.

Reis, Marga. 2005. On the syntax of so-called focus particles in German – a reply to Büring and Hartmann 2001. *Natural Language & Linguistic Theory* 23: 459-483.

Reuland, Eric. 1983. Governing *–ing*. *Linguistic Inquiry* 14: 101-136.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, ed. by Liliane Haegeman, 281-337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 1999. On the position "int(errogative)" in the left periphery of the clause. Ms. Università di Siena.

Rizzi, Luigi. 2004. Locality and the left periphery. In *Structures and beyond: the cartography of syntactic structures*, vol. 3, ed. by Adriana Belletti, 223-251. New York: Oxford University Press.

Rizzi, Luigi. 2006. On the form of chains: criterial positions and ECP effects. In *Wh-movement: moving on*, ed. by Lisa L.-S. Cheng and Norbert Corver, 97-133. Cambridge, Mass.: MIT Press.

Rizzi, Luigi, ed. 2004. *The cartography of syntactic structures*. Vol. 2, *The structure of CP and IP*. Oxford: Oxford University Press.

Romero, Maribel and Chung-hye Han. 2004. On negative *yes/no* questions. *Linguistics and Philosophy* 27: 609-658.

Rooth, Mats. 1985. Association with focus. PhD dissertation. University of Massachusetts.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1: 75-116.

Rothstein, Susan. 1991. Heads, projections, and category determination. In *Views on phrase structure*, ed. by Katherine Leffel and Denis Bouchard, 97-112. Dordrecht: Kluwer Academic Publishers.

Rubin, Edward. 1994. Modification: a syntactic analysis and its consequences. PhD dissertation. Cornell University.

Rubin, Edward. 1996. The transparent syntax and semantics of modifiers. *WCCFL 15*: 429-440.

Rubin, Edward. 2003. Determining pair-Merge. *Linguistic Inquiry* 34: 660-668.

Rullmann, Hortze. 1997. Even, polarity, and scope. *Papers in Experimental and Theoretical Linguistics* 4: 40-64.

Saito, Mamoru and Naoki Fukui. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29: 439-474.

Schreiber, Peter A. 1971. Some constraints on the formation of English sentence adverbs. *Linguistic Inquiry* 2: 83-101.

Sheng, Ji-yan. 2006. Semantic analysis of the modal adverb *ke*. *Journal of Social Science of Jiamusi University* 24.6: 42-44.

Shu, Chih-hsiang. 2006. The syntax of high adverbs: overt and covert positions. Ms., Stony Brook University.

Shyu, Shu-ing. 1995. The syntax of focus and topic in Mandarin Chinese. PhD dissertation. University of Southern California.

Shyu, Shu-ing. 2010. Focus interpretation of *zhi* 'only' associated arguments in Mandarin triadic

constructions. *Linguistics* 48: 671-716.

Soh, Hooi-ling. 1998. Object scrambling in Chinese. PhD dissertation. MIT.

Spenser, Andrew. 1991. *Morphological theory: an introduction to word structure in generative grammar.* Oxford: Blackwell Publishers.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19: 425-449.

Sportiche, Dominique. 1994. Adjuncts and adjunction. *GLOW Newsletter* 32, 54-55.

Sportiche, Dominique. 1998. *Partitions and atoms of clause structure: subjects, agreement, case, and clitics*. London: Routledge.

Stowell, Tim. 2004. Tense and modals. In *The syntax of time*, ed. by Jacqueline Guéron and Jacqueline Lecarme, 621-635. Cambridge, Mass.: MIT Press.

Stump, Gregory. 1981. The interpretation of frequency adjectives. *Linguistics and Philosophy* 5: 221-256.

Sung, Guo-ming. 2007. ji-ge changyong fuci de yuyi ji yuqi [The semantics and pragmatics of some common adverbs]. Paper read at *the 15th IACL and the 19th NACCL Joint Conference*, Columbia University.

Svenonius, Peter. 2002. Subject position and the placement of adverbials. In *Subjects, expletives, and the EPP*, ed. by Peter Svenonius, 201-242. New York: Oxford University Press.

Szabolcsi, Anna. 2006. Strong vs. weak islands. In *The Blackwell companion to syntax, vol. 4*, ed. by Martin Everaert and Henk van Riemsdijk, 479-531. Malden, Mass.: Blackwell Publishing Ltd.

Taglicht, Josef. 1984. *Message and emphasis: on focus and scope in English*. London: Longman.

Taglicht, Josef. 2001. Actually, there's more to it than meets the eye. *English Language and Linguistics* 5.1: 1-16.

Tamori, Ikuhiro. 1979. A study of Japanese adverbs. PhD dissertation. USC.

Tancredi, Christopher. 1990a. Not only *even* but even *only*. Ms., MIT.

Tancredi, Christopher. 1990b. Syntactic association with focus. In *Proceedings from the First Meetings of the Formal Linguistic Society of Mid-America*, 289-303. University of Wisconsin-Madison.

Tang, Jane C-C. 2008. Specifiers vs. non-specifiers: evidence from adjuncts in Formosan languages, Chinese and English. Paper presented at *The past meets the present: a dialogue between historical linguistics and theoretical linguistics*. Academia Sinica, Taipei.

Tang, Sze-wing. 2002. Focus and *dak* in Cantonese. *Journal of Chinese Linguistics* 30: 266-309.

Tang, Ting-chi. 1992. *Hanyu cifa jufa sanji [Studies on Chinese morphology and syntax: 3]*. Taipei: Student Books.

Tenny, L. Carol. 2000. Core events and adverbial modification. In *Events as grammatical objects*,

ed. by Carol Tenny and James Pustejovsky, 285-334. Stanford.: CSLI Publications.

Thomason, Richmond, and Robert Stalnaker. 1973. A semantic theory of adverbs. *Linguistic Inquiry* 4: 195-220.

Thráinsson, Höskuldur. 2000. Object shift and scrambling. In *The handbook of contemporary syntactic theory*, ed. by Mark Baltin and Chris Collins, 148-202. Malden, Mass.: Blackwell Publishing Ltd.

Toman, Jindřich. 1986. A (word-)syntax for participles. *Linguistische Berichte* 105: 367-408.

Toman, Jindřich. 1998. Word syntax. In *The handbook of morphology*, ed. by Andrew Spencer and Arnold M. Zwicky, 306-321. Oxford: Blackwell Publishers Ltd.

Travis, Lisa. 1988. The syntax of adverbs. In *McGill working papers in linguistics: special issue on comparative German syntax*, 280-310. McGill University, Montreal.

Tsai, W.-T. Dylan. 1994. On economizing A-bar dependencies. PhD dissertation. MIT.

Tsai, W.-T. Dylan. 2008a. Left periphery and *how-why* alternations. *Journal of East Asian Linguistics* 17: 83-115.

Tsai, W.-T. Dylan. 2008b. Object specificity in Chinese: a view from the *v*P periphery. *The Linguistic Review* 25: 479-502.

Tsao, Feng-fu. 1988. Topics and clause connectives in Chinese. *Bulletin of the Institute of History and Philology* 59.3: 696–737.

Ura, Hiroyuki. 2000. Case. In *The handbook of contemporary syntactic theory*, ed. by Mark Baltin and Chris Collins, 334-373. Malden, Mass.: Blackwell Publishing Ltd.

Van Craenenbroek, Jeroen. 2004. Ellipsis in Dutch dialects. PhD dissertation. Leiden University.

Van Gelderen, Elly. 2000. The absence of verb-movement and the role of C: some negative constructions in Shakespeare. *Studia Linguistica* 54: 412-423.

von Fintel, Kai and Sabine Iatridou. 2003. Epistemic containment. *Linguistic Inquiry* 34: 173-198.

Vikner, Sten. 1995. *Verb movement and expletive subjects in the German languages*. Oxford: Oxford University Press.

Vries, Mark de. 2005. Coordination and syntactic hierarchy. *Studia Linguistica* 59: 83-105.

Wagner, Michael. 2006. Association by movement: evidence from NPI-licensing. *Natural Language Semantics*. 14: 297-324.

Wagner, Michael. 2009. Focus, topic, and word order: a compositional view. In *Alternatives to cartography*, ed. by Jeroen van Cranenbroeck, 53-86. Berlin: Mouton de Gruyter.

Watanabe, Arika. 2004. The genesis of negative concord: Syntax and morphology of negative doubling. *Linguistic Inquiry* 35: 559-612.

William, Edwin. 1975. Small clause in English. In *Syntax and Semantics* 4, ed. by John P. Kimball, 217-245. New York: Seminar Press.

Williams, Edwin. 1994. A reinterpretation of evidence for verb movement in French. In *Verb movement*, ed. by David Lightfoot and Nobert Hornstein, 189-205. Cambridge: Cambridge University Press.

William, Edwin. 2000. Adjunct modification. *Rivista di Linguistica*. 12: 129-154.

William, Edwin. 2009. There is no alternative to cartography. In *Alternatives to cartography*, ed. by Jeroen van Cranenbroeck, 361-373. Berlin: Mouton de Gruyter.

Wu, Ming-jing. 2009. Yuqi fuzi zai ju zhong weizhi fenbu de yanjiu [A study of the syntactic distributions of mood adverbs]. *Jian Nan Wenxue* 10: 112-114.

Wurmbrand, Susi. 2008. Nor: neither disjunction nor paradox. *Linguistic Inquiry* 39: 511-522.

Yim, Changguk. 2003. Subject agreement in Korean: Move F, Attract F, or Agree? In *TLS 5 Proceedings,* ed. by William Earl Griffin, 147–56. Austin: Texas Linguistics Forum.

Yuan, Yu-lin. 2002. Duoxiang fuci gongxian de yuxu yuanze ji qi renzhi jieshi [The principles of the order of adverbs and their conceptual explanations]. *Yuyanxue Luncong* 26: 313-339.

Zanuttini, Raffaella. 1991. Syntactic properties of sentential negation: a comparative study of Romance languages. PhD dissertation, University of Pennsylvania.

Zeijlstra, Hedde. 2004. Sentential negation and negative concord. PhD dissertation. University of Amsterdam.

Zhang, Nina Ning. 1997. Syntactic dependencies in Mandarin Chinese. PhD dissertation. University of Toronto.

Zhang, Nina Ning. 2008. Repetitive and correlative coordinators as focus particles parasitic on coordinators. *SKY Journal of Linguistics*. 21: 295-342.

Zhang, Yi-sheng. 2000. *Xiandai han yu fuci yanju* [*A study of adverbs in Modern Chinese*]. Shanghai: Xuelin Publishing.

Zhang, Yi-sheng. 2004. *Xiandai hanyu fuci tansuo* [*Explorations of adverbs in Modern Chinese*]. Shanghai: Xuelin Publishing.

Zimmermann, Malte. 2003. Pluractionality and complex quantifier formation. *Natural Language Semantics* 11: 249-287.

Zwart, Jan-Wouter. 2009. Uncharted territory? Towards a non-cartographic account of Germanic syntax. In *Advances in comparative Germanic syntax*, ed. by Alexiadou, Artemis, Jorge Hankamer, Thomas McFadden, Justin Nuger and Florian Schäfer, 59–84. Amsterdam: John Benjamins.

Zwicky, Arnold M. 1977. *On clitics*. Bloomington: Indiana University Linguistics Club.