This is an author-generated prepublished draft. The final publication is available at http://www.springerlink.com (DOI 10.1007/s11049-011-9164-y).

On the Grammar of Extraction and Coordination*

Rui P. Chaves

Linguistics Department – University at Buffalo
The State University of New York

rchaves@buffalo.edu

Abstract

This paper examines English unbounded dependency constructions and concludes that the syntactic mechanism responsible for extraction operates in the same way in all constructions, regardless of the structures being subordinate or coordinate. In other words, this work argues that there is a general and fully uniform mechanism behind the linkage of 'gaps' and their 'fillers' in the family of constructions that allow such unbounded dependencies. Crucially, this mechanism is devoid of exceptions, and interacts with other aspects outside the grammar of extraction to predict the Coordinate Structure Constraint, ATB extractions, parasitism, and at least certain islands. Although the underlying syntactic mechanism for extraction is argued to work in the same way for all constructions, certain parts of the grammar can locally impose additional conditions on the possible extraction patterns. Along with recent psycholinguistic and neurolinguistic research, it is argued that many apparent exceptions result from independently motivated performance factors, also observed in garden-path sentences and center-embedded constructions.

1 Introduction

This work argues that a simple and uniform theory of extraction can be obtained if the datasets that have been assumed for the last half century are properly reassessed. Unlike most previous studies, the empirical evidence considered in this work suggests that the syntactic conditions that govern extraction phenomena in English coordinate and subordinate constructions are essentially the same. Accordingly, I propose a unique general mechanism for the propagation and linking of unbounded dependencies in all constructions, which need not make any distinction between extraction in coordinate structures (symmetric or asymmetric) and extraction in non-coordinate structures.

Levine and Hukari (2006) show in their extensive study that extraction phenomena exhibit various systematic and uniform patterns which have not been generally

*I am very grateful to Ivan Sag and Jean-Pierre Koenig for discussion about an earlier draft. I also thank three anonymous reviewers and the editor Marcel den Dikken for their many invaluable comments. I alone am responsible for the views reflected here and for any errors that remain.

recognized. They provide convincing cross-linguistic evidence that there is a single connectivity mechanism involved in linking fillers and gaps, and point out that many previous proposals are based on non-representative data, and wrongly assume that islands are due to grammatical factors. Various kinds of islands – and their exceptions – may be due to pragmatic and performance effects that are also observed at work in non-extraction phenomena (see Abrusán (2008) and Sag et al. (2007) for overviews). On the other hand, Sag (2010a) argues that the various constructions that exhibit unbounded dependencies are best seen as a family of constructions, with various levels of generalization. Along the lines of the above research, I propose a unified account of extraction for English which operates in exactly the same way for coordinate and subordinate constructions, and predicts the special extraction patterns observed in symmetric coordinate constructions from independently motivated mechanisms. I argue that certain island constraints and parasitism phenomena are the result of independently motivated extra-grammatical factors, and that the coordinate structure constraint and the ATB exceptions follow from semantic-pragmatic factors. Other aspects of coordinate extraction result from an independently motivated non-headed analysis of coordination.

The paper is structured as follows. In §2 I provide a general pre-theoretical overview of extraction patterns, with the goal of showing that the overall range of extraction possibilities is relatively seamless. My base hypothesis is that the underlying syntactic mechanism responsible for extraction is exactly the same across different constructions, coordinate or otherwise. In §3 I review the literature on performance and extraction, and argue along with other researchers that gradient and apparently exceptional extraction patterns are best seen as the result of non-syntactic factors also at work in other phenomena, such as center-embedding and garden-paths. In §4 I provide a brief overview of previous syntactic theories of extraction in coordination. In §5 I propose a syntactic mechanism for extraction in Head-Driven Phrase Structure Grammar (Pollard and Sag, 1994) that both simplifies and generalizes previous accounts. This includes a grammar fragment that can handle the syntax and semantics of different kinds of coordination in a uniform manner.

2 General extraction patterns

In this paper I refer to the missing element as the *gap*, and to the overt element linked to that gap as the *filler*. The link that is established between the two is an *unbounded dependency*. Levine and Hukari (2006) provide an extensive overview of extraction and offer empirical objections against the distinction between parasitic gaps and regular gaps.[1] I subscribe to this view, and will treat all gaps equally. See §2.3 and §3 for more empirical arguments against the notion of parasitism.

[1]For example, Cinque (1990) and Postal (1993,1994,1998,2001) propose two different kinds of extraction. One kind is a canonical *wh*-extraction, and the other corresponds to the insertion of a null resumptive pronoun at the extraction site. In both accounts, the latter corresponds to parasitic extraction. However, Levine (2001) and Levine and Hukari (2006, 256) argue convincingly that the distinction between A/B-extractions is problematic.

2.1 Single extraction

It has been known at least since Ross (1967) that many constructions allow an embedded constituent to be realized in an arbitrarily high position in syntactic structure. These include clefts, topicalization, non-subject relatives, and *wh*-questions. Typically, the extracted element is a subject or a complement, as in (1) and (2), but it is also well-known that certain adverbial structures can be extracted, as illustrated in (3).

(1) a. Who do you think [_ left the party in a hurry]?
    b. This is who I think [_ will win in the documentary category].
    c. Which movie did you think I said [_ would just be a lame parody of Star Wars]?

(2) a. This is the movie that [I really like _].
    b. That actor, I think [I've never seen _ before].
    c. This book will not be easy [to convince young children to read _].
    d. What did you write [a book about _]?
    e. What are you [a doctor of _]?
    f. Which movie did you hear [rumors [that we had boycotted _]]?
    g. Kim is the sort of person that I just don't know [a lot of [people who think well of _]].

(3) a. [How often] do you think that [Fred was late this week _]?
    b. [On Monday], I think that [Kim went home very late _].
    c. [Yesterday], it seems that [Kim arrived home very early _].

Extraction from subject phrases is difficult, but not impossible, as argued by Ross (1967, 242), Pollard and Sag (1994, 195,ft.32), Huddleston et al. (2002, 1093,1094, ft.27), Levine and Sag (2003, 252, ft.6), Kluender (1998, 2004), Levine and Hukari (2006, 71), and others. This is illustrated in (4), for a wide range of constructions.

The examples (4e–i) are my own. These require a prosodic break at the gap site and some contrastive stress on the nominal head of the subject.

(4) a. Of which cars were [the hoods _] damaged by the explosion?
    b. They have eight children [of whom] I think [[five _] are still living at home].
    c. What were [pictures of _] seen around the globe?
    d. Who does [being able to bake ginger cookies for _] give her great pleasure?
    e. Which president would [the impeachment of _] cause outrage?
    f. Which book will [the author of _] never be known?
    g. Which problem will [no solution to _] ever be found?

h. Which crime will [the punishment for _] never be carried out?

i. There are people in this world that [(for me) to describe _ as despicable] would be an understatement.

Not only is it possible to extract from complement phrases and subject phrases, but it is also possible to extract from adjunct phrases. Ross (1967), Chomsky (1982), Engdahl (1983), Hegarty (1990), Cinque (1990), Pollard and Sag (1994), and Borgonovo and Neeleman (2000) among others note that (non-parasitic) extraction from adjuncts is possible, as illustrated in (5).

(5) a. That's the symphony that Schubert [died without finishing _].

b. Which report did Kim [go to lunch without reading _]?

c. A problem this important, I could never [go home without solving _ first].

d. What did he [fall asleep complaining about _]?

e. What did John [drive Mary crazy trying to fix _]?

f. Who did you [go to Girona in order to meet _]?

g. Who would you rather [sing with _]?

I note that this includes some tensed adjuncts, contrary to commonly assumed:

(6) a. Which email account would you be in trouble if someone broke into _?

b. Which problem would you be devastated if someone had already solved _?

c. This is the formula that I would be devastated if someone had already discovered _.

The data above suggest that the extraction mechanism operates in a uniform way across constructions. Various syntactic, semantic, pragmatic, and performance limitations to extraction are discussed in §3. However, Ross (1967) noted that matters are different in coordinate structures since conjuncts can never be extracted. Compare (7) with (8).[2]

[2]Johannessen (1998,216) and Munn (1993,15) argue that (i) and (ii) show that conjuncts can in fact be moved. I disagree with this assessment, and suspect that these data are elliptical verbal coordinations rather than true NP coordinations (for example, instances of *Stripping* (Ross, 1967; Hankamer and Sag, 1976; Chao, 1988)). If these were truly NP conjunctions then (iii) and (iv) should be grammatical, barring further stipulations.

i. All the heaviness had gone, and the height.

ii. John bought a book yesterday, and a newspaper.

iii. *[One man _] agreed with each other [and one woman].

iv. *Both [a man _] smiled [and a woman].

The oddness of (iii) stems from *each other* not having a plural NP antecedent (which conjunction usually forms), and (iv) is odd because expressions like *both* and *between* must combine with a plural NP. Compare (iii) and (iv) with Ross (1967, 14). Other possible cases of conjunct extraction such as *betrayed you he has, and been handsomely rewarded as a result* – brought to my attention by an anonymous reviewer – may result from the interaction of linearization and ellipsis phenomena along the lines of Kathol (1993,1999). I suspect such patterns do not arise in NP coordination because English verbal domains are simply more flexible with regard to word order than the nominal domain.

(7) a. *Who did you see [_ and Tim]?

b. *Who did you talk to [Kim and _]?

(8) a. Who did you [introduce _ to Kim]?

b. What did you talk [to Kim about _]?

Ross (1967) also observed that extraction from inside a conjunct is also prohibited, as shown in the contrast between (9) and (10).

(9) a. *What did Kim [cook _ and wash the dishes]?

b. *What did Kim [cook supper and wash _]?

(10) a. What did Kim [cook _ while watching TV]?

b. What did Kim [cook the rice with _]?

Such contrasts are not a problem for the hypothesis that the syntactic extraction mechanism operates in the same way in all constructions. As many authors have pointed out, the extraction limitations in (9) may be the result of semantic/pragmatic conditions rather than due to special syntactic operations. I discuss this matter in more detail in §4.3.

Still, matters are further complicated by the fact that certain coordinate structures do in fact allow extraction from one conjunct. Consider for instance the data in (11), taken from Ross (1967, 93–94), Schmerling (1972), Goldsmith (1985), and Lakoff (1986). Semantically, these coordinations are different from the cases in (9), which is connected to the different extraction possibilities.

(11) a. Here's the whiskey which I went to the store and bought _.

b. Who did Lizzie Borden take an ax and whack _ to death?

c. How much can you drink _ and still stay sober?

d. How many lakes can we destroy _ and not arouse public antipathy?

Following Levin and Prince (1986), and others, I will distinguish between *symmetric* and *asymmetric* readings of coordinate structures. The two kinds are seen in (12) and (13), respectively. In the preferential reading of (12) the conjunct order is not relevant for semantic interpretation. This contrasts with the preferential readings of the asymmetric coordinations in (13). Here, the order of conjuncts cannot be reversed without semantic contrast, and the conjunction can be paraphrased as *and therefore* or as *and then*. Crucially, the coordinations in (9) are symmetric, while the ones in (11) are asymmetric.[3]

(12) a. Fred likes London and Mia likes Prague.

[3]Typically conjunctions are ambiguous between symmetric and asymmetric readings and context can resolve their interpretations one way or the other. This is particularly clear in CP coordinations, which usually have a tendency for symmetric interpretations, but not always: *Tom said that the gun had gone off and that the bullet hit the dog.*

b. Tom fixed the door and Mary painted the window frame.

(13) a. I dialed 911 and an ambulance arrived.

b. He jumped on his horse and rode off into the sunset.

Ross (1967, 93), and many others since then, assume that asymmetric coordination results from an underlying subordinate syntactic structure, which would in turn explain the extraction differences. However, the non-coordinate analysis for asymmetric readings is problematic. First, Schmerling (1972) notes that the subordinate counterpart of the coordinate structure is in general not synonymous with the coordinate realization. For example, Ross proposed that *I went to the store and bought some whiskey* is synonymous with *I went to the store to buy some whiskey.* Schmerling points out that the two structures are not synonymous, as seen in the contrast in (14). If (14b) is to be derived from (14a) then the meaning difference is unexpected.

(14) a. I went to the store to buy some whiskey, but the sales clerk persuaded me to buy Ripple instead.

b.*I went to the store and bought some whiskey, but the sales clerk persuaded me to buy Ripple instead.

Second, Lakoff (1986) observes that asymmetric constructions behave very much like coordination in the sense that they can be iterated without an upper limit on the number of conjuncts. The data pattern with coordination rather than with adjunction:

(15) a. How many kinds of tequila has he [snuck off to Mexico, sampled _, and come back the same day without telling anyone]?

b. Concerts that short, you can leave work early, hear the entirety of _, and still be back at the job before anyone notices you are gone.

A third aspect that undermines the idea that such asymmetric readings result from an underlying subordinate structure is that coordinate structures in general lack the mobility observed in subordination, as noted by Levine (2001) and Kehler (2002) with data like (16).

(16) a. We can expect our graduate students to teach one course, and still finish a dissertation on time.

b.*And still finish a dissertation on time, we can expect our graduate students to teach one course.

Fourth, in both symmetric and asymmetric coordination none of the conjuncts can be said to be *interpreted* in any intuitive way as a head that governs the realization of the remaining conjuncts. In sum, there is no compelling independent evidence for assuming that symmetric and asymmetric coordination have different syntactic structures.

It is important to note that there are constructions which are similar to coordination – in the sense that none of the daughters is an obvious candidate for being the head – and which behave like subordinate structures and asymmetric coordinate

structures with respect to extraction. Take for example the comparatives shown in (17).[4] The example in (17a) is attributed to Joan Bresnan in Hendriks (1995).

(17) a. Which actor do [[as many men admire _] as [there are women who detest Sylvester Stallone]]?

b. This is the company that they said [[_ hired more consultants] than [PARC hired programmers]].

c. This is the kind of problem$_i$ that [[the sooner you solve _$_i$], [the quicker you'll be on your way to promotion]].

d. [The more theorems]$_i$ we find [someone who can prove _$_i$], the better.

e. This is the kind of scandal$_i$ that [[the richer you are], [the easier it is for you to survive _$_i$]].

Sag (2000) and Bouma et al. (2001) propose that the impossibility of displacing conjuncts may be a consequence of extraction not creating traces, along the lines of Sag and Fodor (1994).[5] In this view, there is simply nothing to conjoin in (7), and thus nothing else needs to be said about conjunct extraction. Sag (2000) revisits various arguments for the existence of traces in the literature and argues that none is unproblematic. While I do not claim that all arguments for the existence of traces are invalid (in fact, some HPSG accounts of extraction like Levine (2003) put traces to good use), I share with Sag (2010a) the serious misgivings about many of the arguments taken for granted in the transformational literature. In this paper I assume a traceless analysis, although my account can be revised to allow the presence of traces.

There is a second possible explanation for the non-existence of conjunct extraction. If coordinators are markers rather than heads, the fact that conjuncts cannot be extracted becomes less remarkable. Still, there is some controversy in the literature about this matter. Although coordinators are not traditionally viewed as heads (e.g. Bloomfield (1933), Yngve (1960), Ross (1967), Pesetsky (1982), Gazdar et al. (1985), Steedman (1989) *inter alia*), Munn (1993), Kayne (1994), and Johannessen (1998) propose that coordinate structures are in fact headed by the coordinator particle. Munn (1993) assumes that coordinators are heads on purely X-bar theoretical grounds, but Kayne (1994) defends this claim with an empirical argument, by claiming that the position of the coordinator is correlated with the position of the verbal head. In verb-final languages the coordinators *tend* to be prospositional while in verb-initial languages the coordinators *tend* to be prepositional. This argument is empirically refuted by the typological data. For example, if the distribution of the coordinator were indeed correlated with the distribution of the verb then one would expect that in freer word order languages the position of the conjunct would be relatively free also. But this does not occur. For example, in Russian the coordinator *i* is required to immediately precede the last conjunct, whereas the verb can be realized in many different

[4]Some comparative constructions have different extraction patterns, as originally noted by Ross (1967,53) and Hankamer (1974). These allow the comparative expression to be stranded, as illustrated by *who is John as tall as _ ?* and *who did you say she arrived earlier than _ ?* For more on comparatives see Borsley (2004) and Sag (2010a).

[5]But see also Abels (2003) and Levine (2003).

positions. More recently, Zwart (2005) presents a survey of 136 languages where half are verb-final and half are verb-initial languages, and compares these according to their initial conjunctions [*conj* NP] and final conjunction [NP *conj*] strategies. The conclusion is that verb-final languages overwhelmingly employ initial conjunction strategies. For example, 119 of these languages have exclusively initial conjunctions, 12 languages exhibit both initial and final conjunctions, and only 4 have exclusively final conjunctions. In sum, there is no correlation between verb position and coordinator position. The conclusions about headedness in Johannessen (1998, 96) are equally problematic. Out of ten criteria for headedness, five are deemed inconclusive and the remaining five are, in my view, invalid (among which is the same argument given in Kayne (1994)). Several arguments are flawed because they ignore the role of semantics. For instance, it is claimed that the conjuncts and the coordinate structure do not have the same syntactic distribution, given examples like *Ruth and Ursula embraced* and **Ruth embraced*. From this Johannessen concludes that the conjunction lexeme *could* be the head of the structure. However, this oddness is clearly not due to syntax, but to semantics: the verb *embrace* requires a pluralic subject (cf. with *Two boys* (*and two girls*) *embraced*). Thus, the conjuncts and the coordinate structure *do* have the same syntactic distribution once one takes into account semantics.[6] In sum, there are good syntactic reasons for viewing coordinate structures as non-headed. For further controversial aspects see Borsley (2005).

In conclusion, the evidence suggests that one can in principle extract from essentially any kind of daughter, in a variety of different constructions. In particular, we have seen that headed and non-headed constructions exhibit essentially the same extraction patterns. An independently motivated traceless and non-headed analysis explains why conjuncts cannot be extracted, and parallelism conditions active in symmetric coordination (arguably of non-syntactic nature) constrain the extraction mechanism. These observations are consistent with the idea that all constructions (headed or non-headed, coordinate or otherwise) are subject to the same *syntactic* extraction mechanism.

2.2 Multiple Extraction with a common Filler

It is well-known that gaps originating in different daughters can be 'fused' in the sense that they are semantically linked to the same filler. Again, this phenomenon can occur in various kinds of headed constructions, as (18) shows. For more discussion about examples like (18e) see Kayne (1983).

[6]Arguments about 'unbalanced coordination' – where one of the conjuncts has privileged features that must be percolated to the mother node – are unconvincing because the data are fully compatible with a non-headed analysis. For example, Korean *ko/kwa* conjunction is remarkable because only like categories are conjoinable but case and mood is only recorded in and percolated from the final conjunct. However, the fact that one conjunct seems to function as the head with regard to two features but not to others does not entail that it is a syntactic head (and it clearly is not a semantic head). In the rule below I illustrate how to capture these facts in a construction-based approach. Both conjuncts (the second of which is assumed to be marked by a coordinator) have the same pos and valence information as the mother, but only the second conjunct has mood and case to share: {pos:X,valence:Y,case:W,mood:Z} → {pos:X,valence:Y,case:none,mood:none} {pos:X,valence:Y,case:W,mood:Z}.

(18) a. Who$_i$ did you [give [pictures of _$_i$] [to _$_i$]]?

b. That was the rebel leader who$_i$ [[rivals of _$_i$] shot _$_i$].

c. Here's the guy$_i$ that I [expected [my pictures of _$_i$] [to bother _$_i$]].

d. Which model$_i$ did you say [Mary [fell in love with _$_i$ [while she was photographing _$_i$]]?

e. She is the kind of person$_i$ that [[everyone who meets _$_i$] ends up falling in love with _$_i$]].

The data point in (19), due to Levine and Sag (2003), shows that simultaneous extraction from a subject and from an adjunct phrase is also possible.

(19) What kinds of books do authors of _ argue about royalties after writing _?

Crucially, multiple extraction with a common filler also occurs both in symmetric and in asymmetric coordination. This is illustrated in (20) and (21), respectively. As is well-known since Ross (1967) and Goldsmith (1985), symmetric coordination requires all conjuncts to contain a gap (Across-the-board extraction, henceforth ATB) while asymmetric coordination does not (compare (9) and (20), for example).

(20) What$_i$ did Kim [cook _$_i$ for two hours and eat _$_i$ in four minutes]?

(21) a. What$_i$ did he go to the store, buy _$_i$, load _$_i$ in his car, and unload _$_i$?

b. How many courses$_i$ can you take _$_i$ for credit, still remain sane, and still get all A's in _$_i$?

Since subjects and complements can be extracted, it follows that a filler can be linked to both a subject gap and a complement gap, as borne out from the data below. The examples in (22a) are from Goodall (1987), Levine et al. (2001), and Williams (1978). Apparent exceptions are discussed in §3.

(22) a. We went to see a movie which [the critics praised _$_i$] but [_$_i$ was too violent for my taste.]

b. Robin is the only person who [_$_i$ likes me] and [I like back _$_i$].

c. I know someone [who Bill has met _$_i$] and [who I think _$_i$ might like Mary].

The data indicate that multiple gaps can be fused and linked to one and the same filler in coordinate and non-coordinate structures alike. Thus, it is plausible that the extraction mechanism operates fundamentally in the same way in all of these constructions.

2.3 Multiple Extraction with multiple Fillers

Filler-gap linkages cause a significant cognitive burden on the language processor for various reasons. Identifying the gap is not easy because it has no overt manifestation and its location must be inferred. A filler must be kept in working memory, while all other material between the filler and the gap is processed. Moreover, fillers remain in working memory even after they are linked to a gap, given that there may be more gaps further downstream as in the data in §2.2. Cases with multiple gaps and fillers are particularly taxing because the correct filler must be linked to the right gap. Although multiple extraction with different fillers is cognitively demanding, such cases exist as illustrated by the data in (23), due to Deane (1992), Fodor (1992), Pollard and Sag (1994, 169), Levine and Hukari (2006, ch.5).

(23) a. [A violin this well crafted]$_i$, even [the most difficult sonata]$_j$ will be easy to play _$_j$ on _$_i$.

b. This is the person who$_i$ I can't remember [which papers]$_j$ I sent copies of _$_j$ to _$_i$.

c. [Someone that stupid]$_i$, [how much time]$_j$ [do we really want to waste _$_j$ arguing with _$_i$]?

d. This is the evidence$_i$ that we need to find someone [who$_j$ we can intimidate _$_j$ with _$_i$].

e. Robin is someone who$_j$ I never know what$_i$ [to say _$_i$ to _$_j$].

In (23a,b), a head is combining with two complements. Each of them contains a gap, and each gap is linked to a different filler. However, in (23c) it seems that one gap is in a head phrase and another gap is in an adjunct phrase. This sort of example raises further questions about the notion of parasitism (see also Levine et al. (2001, 185, ft.7)). Consider the data in (24), validated by native speakers. These show that a gap located in what is clearly a modifier phrase does not need to be linked to a gap located in the head phrase.

(24) a. [A problem this complex]$_i$, [how much time]$_j$ could they spend _$_j$ before solving _$_i$?

b. This was [the kind of person]$_i$ that even [the simplest problem]$_j$ became difficult to solve _$_i$ without shouting at _$_j$.

c. He is [the kind of assistant]$_i$ that even [the *simplest task*]$_j$ is impossible to undertake _$_j$ *without* first escorting _$_i$ outside of the building.

The data discussed so far indicate that any phrasal daughter – be it a head or not – can in principle contain a gap, and that the fusion of multiple gaps is optional as far as syntax is concerned. As we shall see, in certain conditions gap fusion is preferential and in others it is not.

According to the same informants, asymmetric coordinations also allow this extraction pattern, as shown in the causal/temporal-precedence conjunctions in (25). The processing difficulties caused by the complexity of these examples can be reduced if the

fillers are prosodically contrasted. This helps the comprehender retrieve the correct filler and link it to the right gap. Note that in (25c) both gaps are embedded in an adjunct phrase.

(25) a. A problem$_i$ this hard, [how many years]$_j$ could someone spend _$_j$ in the field and not even come close to solving _$_i$?

b. A bird$_i$ this unique, [how many years]$_j$ could someone spend _$_j$ and never actually see _$_i$?

c. This is the guard dog$_i$ that even [the smallest leaf]$_j$ could ruin everything [if we rustled _$_j$ and woke _$_i$].

Symmetric coordination imposes a stronger constraint, as Ross (1967) noted, since the gaps in symmetric coordination must be fused and linked to the same filler. Compare (23) and (26).

(26) a.*[A violin this well crafted]$_i$, even [the most difficult sonata]$_j$ will be easy to write _$_j$ and to play it on _$_i$.

b.*[Someone that stupid]$_i$, [how much time]$_j$ [do we really want to waste _$_j$ and to meet _$_i$]?

Note that this does not mean that a unique filler is required by symmetric coordination. In (27) I provide an example where each conjunct has a pair of distinct gaps that are extracted ATB. Rather, symmetric conjunction merely requires that the conjuncts share the same extracted elements.

(27) [A project this complex]$_i$, [how much time]$_j$ [would he [waste _$_j$ working on _$_i$] and [spend _$_j$ complaining about _$_i$]?

Although each head has at most one extracted dependent in the data that we have seen so far, this need not always be the case. Baltin (1982) and Levine and Hukari (2006, 74) convincingly show that under the correct structural conditions two complements of the same head can be extracted, as in (28a–c). The speaker-validated examples that I provide in (28d,e) suggest that any co-valents are in principle extractable.[7]

(28) a. This is a man [to whom]$_i$ liberty$_j$ we would never grant _$_j$ _$_i$.

b. Ten thousand dollars is a sum of money which$_i$ [to a cause like THAT]$_j$ I would GLADLY give _$_i$ _$_j$.

c. [On which violins]$_i$ are [these sonatas]$_j$ difficult to play _$_j$ _$_i$?

d. A function is a kind of relation that$_i$ [for a GIVEN input]$_j$ (we know) [_$_i$ always returns [the SAME output] _$_j$].

e. Kim is someone who$_i$ [to ANY GIVEN TABLE]$_j$, I suspect [_$_i$ always sends [the WRONG order] _$_j$].

[7]The examples in (28d,e) are inspired in Levine and Hukari (2006, 96).

Let us take stock. For coordinate and non-coordinate constructions alike, any daughter can in general contain a gap, and multiple gaps can be linked to either the same or different fillers. Symmetric conjunction imposes stronger conditions which may be the result of non-syntactic conditions. In §4.3 I discuss how semantic and pragmatic factors can independently lead to the peculiar ATB extraction patterns observed in symmetric coordination. As far as syntax is concerned, the extraction mechanism seems to operate in the same way in coordinate and non-coordinate structures.

2.4 Grammatical and extra-grammatical conditions

Postal (1998,2000) considers a large body of cases and shows how heterogeneous island phenomena are: specific constructions constitute islands for certain kinds of syntactic dependencies but not for others. A construction-based theory such as the one adopted in this paper is able to capture the fact that only certain mother-daughter configurations allow extraction. On the other hand, there is mounting evidence that some kinds of island may result from pragmatic factors (e.g. presupposition failures in tenseless wh-islands, factive islands, and negative islands (Kroch, 1989; Oshima, 2007; Abrusán, 2008)) and the cumulative effect of processing difficulties (Deane, 1991; Pritchett, 1991; Kluender, 1992; Kluender and Kutas, 1993; Kluender, 2004; Levine and Hukari, 2006; Sag et al., 2007; Hofmeister, 2007b). If gradient unacceptability can be due to extralinguistic factors, then it is expected that structurally very similar sentences can differ in acceptability. For example, it is traditionally assumed that a *wh*-question involving a gap within a nonfactive *that*-complement as in (29a) is grammatical, but less so when the same configuration involves an indirect question as in (29b). The contrasts in (29) suggest that indirect questions are 'weak' islands but relative clauses are 'strong' islands.

(29) a. Who does John think Mary will choose _ ?

b.?Who did John wonder whether they will fire _ ?

c.*Who did John meet the girl who will marry _ ?

But it is at least plausible that the graded oddness in (29b) is due to low acceptability rather than to low grammaticality. The notion of graded grammaticality is hard to justify, let alone quantify or measure since all sentence judgments are about acceptability. On the other hand, the above islands have well-known grammatical counterexamples. First, compare (29b) and (30).

(30) a. Who did John wonder whether or not they should fire _ ?

b. Which shoes are you wondering whether you should buy _ ?

Second, virtually all island constraints for NPs that have ever been proposed for cases like (29c) – the A-over-A principle, the Complex NP Constraint, Subjacency, etc. – have known grammatical counterexamples like (31).[8]

[8]Such data were noted by Ross (1967), Kuno (1976, 423), McCawley (1981, 108), Maling and Zaenen (1982), Chung and McCloskey (1983), Fodor (1983), Hegarty (1990), Deane (1991), Pollard and Sag (1994), and Postal (1998).

(31) a. Which astronaut did you read [a book about _]$_{NP}$?

b. Who would you approve of [my seeing _]$_{NP}$?

c. Who did you take [pictures of _]$_{NP}$?

d. Which book did you like [the cover of _]$_{NP}$?

e. Who was the actor that you became [good friends of _]$_{NP}$?

f. This is the paper that we need to find [someone who understands _]$_{NP}$.

g. Violence is something that there are [many Americans who condone _]$_{NP}$.

h. We have a visitor who there's [no one who's willing to host _]$_{NP}$.

i. Which diamond ring did you say there was [nobody in the world who could buy _]$_{NP}$?

j. Which rebel leader did you hear [rumors [that the CIA assassinated _]$_{CP}$]$_{NP}$?

k. It was a new company that Simon spread [the rumor [that they started _]$_{CP}$]$_{NP}$.

l. Nixon was one president that they had [no trouble [getting [votes for the [impeachment of _]$_{NP}$]$_{NP}$]$_{NP}$].

At this point, there are at least two analytical possibilities. Either competence grammar contains numerous fine-grained and theory-internal stipulations about what hampers/facilitates extraction, or competence grammar has little to say about extraction limitations but the acceptability can be hampered by performance-based limitations and pragmatic plausibility. The latter avenue is most promising for two reasons. First, graded examples can often be ameliorated simply by contextualization. Second, the same factors behind gradient and apparently exceptional acceptability in extraction phenomena can be seen at work in other phenomena such as garden-path sentences and center-embedding. I will discuss these aspects in more detail below.

3 The role of performance and coherence constraints

Chomsky (1965) and many others have pointed out that grammatical sentences can be deemed unacceptable by speakers by many reasons, including memory limitations, intonation, stylistic factors, context, and so on. Recently, neurolinguistic and psycholinguistic studies like Kluender (1998), Sag et al. (2007), Hofmeister (2007a,b), Hofmeister and Sag (2010) provide evidence suggesting that degraded extractions may be explained by general performance conditions rather than by competence grammar.

Take for example the classical garden-path *the horse raced past the barn fell*. This is typically deemed unacceptable by native speakers, but the garden-path effect can be reduced. For example, Crain and Steedman (1985) show that the garden-path effect is reduced if context makes clear that there are various horses. The garden-path effect is also reduced with a different choice of words that makes the grammatical parse more likely, as in *the thief arrested by the police turned out to be our nephew*. Pearlmutter and MacDonald (1979) argue that one of the causes of the difficulty in parsing garden-path sentences the to lack of plausibility: the NP *the thief* is not a good subject

phrase for the verb *arrested* because of world knowledge. MacDonald et al. (1994) also point out that *raced* leads to such a garden path because this word form occurs more frequently as a finite verb than as a participle. Other grammatical sentences are perceived as odd for similar reasons. For instance, center-embedding clauses are known to become easier to process (and more acceptable) if the elements are pronominal and bear a coherent and easily recognizable semantic relationship (Powell and Peters, 1973; Stolz, 1976). Hence, (32b) is much better than (32a).

(32) a. [The boy the cat the dog bit scratched] started crying.

b. [The movie everyone I know loved] was *Inception*.

An extreme and well-known example of performance interfering with acceptability is given below. Example (33b) is much harder to parse than the isomorphic example in (33a). Again, the acceptability contrast is due to performance rather than competence. From now on I follow Kluender (1998) in using the symbol '>' to mean *seems better/easier than*.

(33) a. People we love say things we like. >

b. Police police police police police.[9]

Such factors are arguably at work in a wide range of extraction phenomena. Fiengo and Higginbotham (1981), Kluender (1998), and others show that extraction from NPs having definite reference is more complex and less acceptable, as illustrated by the data in (34).

(34) a. Which rebel leader did you hear [rumors that the CIA assassinated _]? >

b. Which rebel leader did you hear [a rumor that the CIA assassinated _]? >

c. Which rebel leader did you hear [the rumor that the CIA assassinated _]? >

d. Which rebel leader did you hear [Tom's rumor that the CIA assassinated _]?

Note that prenominal genitives do not impose any syntactic prohibition to extraction as shown by Kuno (1987, 13) and others:

(35) a. This is the story that I haven't been able to get Mary's version of _.

b. Which theorem did you read Kripke's proof of _?

Kluender (1998) and Hofmeister (2007b) also demonstrate that the extracted element exhibits the opposite tendency: it is easier to process extractions with more specific/informative fillers. Thus, the extraction of a non-specific NP is not as acceptable if the intermediate NPs are specific:

[9]This can be paraphrased as *police policed by police are policed by police*. On the origin of such examples see http://www.cse.buffalo.edu/∼rapaport/BuffaloBuffalo/buffalobuffalo.html.

(36) a. Which article do you need to find someone [who can understand _]? >
b. What do you need to find an expert [who can understand _]? >
c. What do you need to find the professor [who can understand _]?

Erteschik-Shir (1973), Deane (1992) and others argue that the acceptability of extractions can be sensitive to the verbs occurring between the filler and the gap. The examples below involve extraction out of an indefinite NP, which one would expect to be equally good.

(37) a. Who did you see a picture of _? >
b. Who did you destroy a picture of _?

(38) a. Who did you read a book about _? >
b. Who did you tear up a book about _?

Similarly, extraction from objects of experiencer verbs is graded:

(39) a. ?*Who do you usually love a sonata by _?
b. ?*Who do you generally detest an opera by _?

Erteschik-Shir (1981) argues that such acceptability differences stem from pragmatic phenomena. For example, the contrast between the two sentences in (40) arguably stems from the fact that *write a book about Nixon* is more coherent than *destroy a book about Nixon*. Crucially, this approach explains why such extractions can sometimes be ameliorated. For example, if John is known to usually destroy books, then and (40b) becomes fully acceptable. Compare also (36b) with Kuno's *who did they destroy more pictures of, Chairman Mao or Jiang Qing?* Cowart (1997,314) also provides other examples of verb-dependent variability in extraction acceptability, where factive clauses are more permeable to extraction when compared with assertive clauses.

(40) a. Who did John write a book about?
b. ?Who did John destroy a book about?

Why should specificity, distance, and frequency play any role in the *competence* grammar of extraction? The answer given by Kluender (1992, 1998), Fanselow and Frisch (2004), Sag et al. (2007), Hofmeister (2007b) and others is that the grammar of extraction contains no such restrictions. Rather, such islands result from the accumulation of cognitive costs in accessing infrequent lexemes, keeping track of semantically light fillers in memory, and linking fillers to gaps when the intervening structures are semantically complex. While semantically light verbs facilitate extraction, semantically richer verbs hamper it because they consume more memory resources when the sentence is being processed. In sum, the interplay of syntax, processing and pragmatics creates a complex landscape of gradient acceptability judgements.

This is consistent with Alexopoulou and Keller (2007), where it is shown that there is cross-linguistic evidence of performance effects behind gradient acceptability, and

with Sag et al. (2007), where so-called 'D-Linking' effects (Pesetsky, 1987, 2000) follow from more general factors pertaining to memory retrieval/decay. For criticism of D-linking on both empirical and conceptual grounds see Kroch (1989), Ginzburg and Sag (2000, 247–250), Levine and Hukari (2006, 242, 268–271), Hofmeister (2007a) and Hofmeister and Sag (2010). If this is correct, then a proper account of the syntax of extraction should remain silent about the above gradient and apparently exceptional cases. Their markedness and oddness may be due to the interplay of pragmatics and the independently motivated cumulative effect of resource-limitation performance constraints.

3.1 On parasitism effects

Let us consider some possible extragrammatical explanations for parasitism. What could be a plausible explanation for the unacceptability of (41a), and the graded acceptability of (41b)? Although the example in (41c) – due to Levine and Sag (2003) – is not trivial to process, it does not deserve to be labeled as ungrammatical.

(41) a.*Which person did you invite me without thinking _ would actually come?

b.?Which person did you invite _ without thinking I would actually come?

c. Which person did you invite _ without thinking _ would actually come?

Pickering and Barry (1991), Kluender (1998), Chen et al. (2005) and many others show that keeping a displaced item in memory has a cognitive cost, and thus there is a strong tendency to postulate a gap, link it to the filler, and discharge it from memory *as soon as possible*.[10] The first formulation of this performance condition was the 'Active Filler Hypothesis' (Frazier, 1987), illustrated in (42). All things being equal, the sentence *who did you tell Mary left the country?* is more likely to be parsed as (42a) than as (42b).

(42) a. Who did you tell _ Mary left the country?

b. Who did you tell Mary _ left the country?

The same is observed by Stowe (1986), who finds that subjects take significantly longer to process the pronoun *us* in (43c) than in (43a,b). The parser predicts a gap in the object position of *bring* in (43c) and fills it with *who*. This filled gap is contradicted by the presence of the overt pronoun *us*, leading to a reanalysis and longer processing time. No reanalysis occurs in (43a,b).

(43) a. My brother wanted to know if Ruth will bring us home to Mom at Christmas.

b. My brother wanted to know who will bring us home to Mom at Christmas.

c. My brother wanted to know who Ruth will bring us home to _ at Christmas.

[10]This is also consistent with the fact, orginally noted by Fodor (1978), that acceptability degrades in multiple extraction sentences when extraction paths are crossed.

Another kind of example is identified by Ellis (1991), who points out that sentences like (44) exhibit a garden-path because the parser is biased to take *which rumors* as the displaced object of *believes*. Of course, speakers can choose to aid comprehenders by providing prosodic cues, but the point is that there are isomorphic sentences which do not exhibit this garden-path effect. For example – all things being equal – the sentence *which rumors do you think John said offended many people?* is easier to process than (44) because *rumors* is not a good candidate for being the complement of *think*.

(44) Which rumors did you say John believes offended many people?

Moreover, when a verb has several subcategorization frames, but is biased towards a particular one, the parser can be misled when postulating the location of a gap. In (45), for example, there is a very strong tendency to view the verb *buy* as transitive rather than as ditransitive.

(45) What did you say I would never consider buying you _ ?

Thus, the oddness of (41a) may stem from this kind of processing bias, since it is riddled with such minor garden-paths. The first arises when parsing *which person did you invite*, where there is a strong tendency to postulate a complement gap for *invite* and link it to the filler. When the pronoun *me* is encountered, the parser must abandon that parse and backtrack. Another instance of backtracking occurs when parsing *which person did you invite me*, since there is a tendency to try to view *invite* as ditransitive (e.g. ?**who did you invite* (*for*) *us this time?*). A similar effect is discussed in Engdahl (1983) about the tendency to view the intransitive verb *sneeze* as a transitive in *Who did you sneeze after meeting* _ ? The parser backtracks a third time when encountering *which person did you invite me without*, since there is a tendency to postulate a gap as the complement of the preposition. In sum, there are multiple points at which parsing goes awry before the correct filler-gap linkage in (41a) can be identified. On the other hand, there are fewer such points in (41b) and (41c). I will return to these examples shortly.

The acceptability of (46) is graded because of similar performance factors. The garden-path effect can be reduced with a pause after *jogs*.

(46) Since Kim always jogs a mile seems like a short distance to her.

A second major factor that is known to interfere with acceptability is lack of pragmatic/contextual coherence. Consider the data in (47). While (47a) conveys a coherent proposition (someone found a photo not being sought), (47b) does not (someone found a photo without looking for a book). The extra processing load caused by the presence of extraction can only make matters worse for the human processor.

(47) a. That is the photo that I found _ without looking for _ .

b.?That is the photo that I found _ without looking for a book.

The oddness in (47) is not configurational since one can remove the pragmatic interference and ameliorate the same extraction pattern:

(48) a. This is the car that he bought _ without leaving the house.

b. That is the photo that I found _ while looking for a book.

Now consider the examples in (49), noted by Levine and Sag (2003). The low acceptability of (49b,c) can also be seen as the result of the same kind of cumulative effect of pragmatic failure and processing difficulty.

(49) a. What kinds of books do authors of _ argue about royalties after writing _ ?

b. ??What kinds of books do authors of malicious pamphlets argue about royalties after writing _ ?

c. ??What kinds of books do authors of _ argue about royalties after writing malicious pamphlets?

The oddness of (49b,c) may result from both coherence and processing phenomena. In (49b), replacing the highly specific intervening NP *authors of malicious pamphlets* by *they*, and replacing *do* by *did* significantly improves the acceptability of the sentence. In (49c) there is a strong tendency to view the object of *about* as a gap, and the sentence is not fully coherent (it is odd that writing malicious pamphlets would lead to book authors arguing about royalties).

On the other hand, (49a) is coherent, and the existence of multiple co-referential gaps may actually *aid* the language processor by reinforcing the parser's expectations about a filler-gap linkage choice. I conjecture that this is related to a faciliatory process that has been observed by Vasishth and Lewis (2006) and others – known as 'anti-locality' – where sentence processing can be facilitated by *increasing* argument-verb distance with expressions that reinforce expectations about the existence of an upcoming verb. They argue that intervening material can lead to repeated memory retrieval and hence to the reactivation of dependents. I conjecture that the multiple co-referential gaps in (49a) and in (41c) may be faciliatory for the same reason.

A final example of the explanatory role of processing is provided by the contrast in (50). Extraction from subjects is well-known to be particularly difficult, and it is standardly assumed that subjects can only be gapped if their verbal head contains a gap. However, (50b) is not exactly easy to process either, and the non-extracted counterpart of (50a) is stylistically marked to begin with (*my talking to Whitney bothered Hilary*). Hence, (50a) may be odd because of the cumulative effect of processing and pragmatic problems, as argued by Kluender (2004).

(50) a.*Who did my talking to _ bother Hilary?

b. Who did my talking to _ bother _ ?

Evidence in favor of a processing based explanation is provided by the contrast in (51). In this particular example the processor is less likely to be mislead when the gap site is clause-embedded than when it is not clause embedded. These observations are consistent with the hypothesis that subject islands are at least in part due to extragrammatical factors. See Kluender (2004) for more discussion on processing-based explanations for subject islands.

(51) a. ??Which country was [the capital of _] attacked?

b. (?)Which country did you say that [the capital of _] was attacked?

3.2 On unacceptable multiple extractions

There are also cases where multiple extraction is unacceptable. Again, a competence grammar account would have to stipulate complex constraints that shed no light on the nature of these exceptions. Consider (52a) and and (52b) (Williams, 1978; Gazdar, 1981).

(52) a.*I know a man who [[Bill saw _] and [_ liked Mary]].

b. I know someone [who [Bill has met _] and who [I think _ might like Mary]].

I believe that the oddness of (52a) is not due to competence grammar, but rather to preemption caused by multiple alternative parsings that cause the parser to crash. The string *likes Mary* seems to be preferentially parsed as a VP not containing a gap. First, it is seen as a conjunct of the VP *saw*. Backtracking occurs and then *saw and liked Mary* can be viewed as a VP – via Right-Node Raising, for example. For some reason, processing breaks down before the parser can consider the correct subject relative parse for *likes Mary*. Perhaps this is because the verb form is compatible with both the relativized noun and the subject *Bill*, and because of the natural tendency to take the most recent nominal as the subject of both 'seeing' and 'liking'. Conversely, (52b) has many clues leading to the correct parse, and thus is easier to process.

If performance is behind the unacceptability of (52a), then one should be able to ameliorate it by using a verb form that is only compatible with the relativized argument. Consider the data in (53), which are passable with a pause after the first conjunct, without which there is still tendency to parse the second relativized conjunct as a VP rather than as a subject relative clause.

(53) a. There were some parts that [[I enjoyed _] and [_ were very suspenseful]].

b. I ate often, and I ate foods that [[I liked _] and [_ were good for me]].

Any competence account of extraction idiosyncrasies will be hard-pressed to deal with all of the above data. As Kluender (1992), Kluender (1998), Kluender (2004), Fanselow and Frisch (2004), Sag et al. (2007) and others argue at length, a performance account may explain away such idiosyncrasies as the consequence of more general factors (e.g. working memory limitations, frequency, ambiguity management, context, world knowledge, and pragmatic coherence), also measurable when processing garden-path sentences or center-embedded sentences (Warren and Gibson 2002,2005).[11]

4 Accounts of extraction in coordination

Ross (1967, 89) identified a number of conditions governing the displacement in coordination, collectively called the *Coordinate Structure Constraint* (henceforth CSC). The version given in (54) includes the addendum for the ATB exceptions, introduced later in Ross (1967, 96). I will follow Grosu (1973) and refer to (i) as the *Conjunct Constraint* and to (ii) as the *Element Constraint*.

[11]Culicover and Levine (2001) suggest *that*-trace effects are a garden-path phenomenon.

(54) Coordinate Structure Constraint:

in a coordinate structure, (i) no conjunct may be moved, (ii) nor may any element contained in a conjunct be moved out of that conjunct ... unless each conjunct properly contains a gap paired with the same filler.

It is clear that the Conjunct Constraint and the Element Constraint are of very different natures. The former does not have an ATB exception, and operates in both symmetric and asymmetric coordination, as seen in (55). As noted in §2.1, the latter data are unremarkable in a traceless grammar. In what follows, I will briefly discuss the main attempts to explain the CSC and the ATB exceptions.

(55) a. How did you feel _ ?

b. Did you feel betrayed and (then) sad?

c.*How did you feel _ and (then) sad?

d.*How did you feel both _ and _ ?

4.1 Transformational approaches

Capturing the CSC phenomena and the ATB exceptions in a satisfactory way has always posed a major challenge for transformational grammar. The main problem lies in reconciling the kind of movement operation that is needed in coordinate structures with the movement operation needed elsewhere, without creating inconsistencies in the theory or resorting to stipulations that merely restate the CSC. For example, Williams (1978) essentially restated the CSC by resorting to a special kind of ATB-movement mechanism that is only active in coordination. Stipulating exceptional and special-purpose operations should be avoided in favor of a more general mechanism. In transformational grammar, it remains a recalcitrant fact that multiple extraction in (symmetric) coordination entails one filler, but multiple extraction in other structures does not. Another problem concerns stipulating that traces cannot be conjoined, which should be a legitimate option.

In Goodall (1987) coordinate structures are treated in terms of the union of reduced phrase markers. As a result, the ATB exceptions are predicted. However, Goodall notes that this approach cannot account for the Conjunct Constraint. Thus, additional stipulations are needed. Along the lines of the *nested dependency condition* from Fodor (1978), Pesetsky (1982) proposes the *Path Containment Condition* (PCC), which constitutes an attempt to give a generalized theory of constraints on crossing nested dependencies and of the CSC. A path is defined as the set of nodes in a tree structure that connect the head of a chain to the foot of the chain. The PCC states that if two paths overlap, one must contain the other. In this account, one would expect that extractions with multiple gaps and multiple fillers like the ones in §2.3 are impossible. In these cases the paths clearly overlap. Similarly, cases like (27), repeated below as (56), would also be ruled out as impossible.[12]

[12]Johannessen (1998,228–235) assumes that sub-clausal coordination is obtained from clausal coordination, and so it is not clear how the account can cope with *what kind of herbs can I both* [*eat and make tea from*]*?* since the clausal counterpart is impossible: **what kind of herbs can I both eat and can I make tea from?*

(56) [A project this complex]$_i$, [how much time]$_j$ [would he [waste _ $_j$ working on _ $_i$] and [spend _ $_j$ complaining about _ $_i$]?

Postal (1998) proposes the existence of two types of extraction: A-extraction, which leaves behind a trace, and B-extraction, which leaves an invisible resumptive pronoun. The counterexamples to the CSC are seen as either instances of B-extraction, and thus not true counterexamples, or as not involving true conjunction. Levine (2001) and Levine and Hukari (2006,ch.4) offer a detailed assessment of Postal (1998) and provides a number of counterexamples to the data on which Postal's account is based, and argue against the non-coordinate status of asymmetric coordination (cf. with §1.1).

4.2 Non-Transformational approaches

Non-transformational accounts of extraction are typically based on the 'slashed category' analysis of Gazdar (1981). Since there is no notion of structural displacement between trees, the problem created by transformations does not exist, and it becomes trivial to model different gaps linked linked to the same filler without creating a fundamental inconsistency. For exposition purposes, I discuss below in general terms how the CSC facts are obtained using a simplified phrase-structure grammar fragment based on Pollard and Sag (1994) and Sag (2010b), rather than the formalization in Gazdar (1981).

Lexicalist non-transformational theories typically assume that part-of--speech and subcategorization constraints are associated to both lexical and phrasal expressions. This information can be encoded in typed feature structures, as in the lexical entry for the verb *read* in (57). The valence feature VAL(ENCE) contains a list of *in situ* dependents while SLASH contains a set of dependents that are not realized *in situ*, but rather, enter a long-distance dependency. The symbol 'NP' is an abbreviation for the features [*noun*, VAL ⟨⟩], 'VP' is an abbreviation for [*verb*, VAL ⟨X⟩], and 'S' for [*verb*, VAL ⟨⟩].

(57) $$\textit{read} : \begin{bmatrix} \textit{verb} & \\ \textsc{val} & \langle\, \mathrm{NP}_i,\, \mathrm{NP}_j \,\rangle \\ \textsc{slash} & \{\,\} \end{bmatrix}$$

Grammar rules are responsible for allowing alternative realizations of this verb. Some have no major impact on valence (e.g. inflectional rules), while others do (e.g. the passivization rule inflects the verb, and changes the VAL feature from ⟨ NP$_i$, NP$_j$ ⟩ to ⟨ NP$_j$ (, PP$_i$) ⟩). A similar lexical process can be argued to license extraction: a rule can take a verb like (57) and produce (58), where the NP complement will not be realized *in situ*. The valent will not be recorded in VAL, but rather, in SLASH.

(58) $$\textit{read} : \begin{bmatrix} \textit{verb} & \\ \textsc{val} & \langle\, \mathrm{NP}_i \,\rangle \\ \textsc{slash} & \{\, \mathrm{NP}_j \,\} \end{bmatrix}$$

Although in some cases both variants of the word exist, one with gaps and another without, in other cases only the gapless form is allowed in the grammar. For example,

in languages that do not allow preposition stranding the realization of prepositions with their complement in SLASH is not permitted. Conversely, in other instances only the non-empty SLASH counterpart exists, as in the case of verbs like *assure*. These must have one complement in SLASH rather than in VAL, as suggested by the examples in (59), based on Kayne (1980). Finally, in the case of *tough*-movement words, the selected complement must have a non-empty SLASH specification as discussed above in §2.4,

(59) a.*I can assure you him to be the most competent.

b. Who$_i$ can you assure me _ $_i$ to be the most competent?

The rules that deal with the realization of *in situ* valents are given in (60), in simplified form. The first rule allows a head of any category X that selects complements $Y_2 ... Y_n$ to combine with such complements and yield a constituent subcategorizing for a specifier Y_1. The second rule allows the latter category to combine with the subcategorized specifier. An NP is thus a nominal category with the feature [VAL ⟨⟩], a VP is a verbal category with the feature [VAL ⟨NP⟩], a sentence is a verbal category with [VAL ⟨⟩], and so on.

(60) a. $X_{\text{VAL } \langle Y_1 \rangle} \rightarrow X_{\text{VAL } \langle Y_1, Y_2, ..., Y_n \rangle} \quad Y_2 \ldots Y_n$

b. $X_{\text{VAL } \langle \, \rangle} \rightarrow Y_1 \quad X_{\text{VAL } \langle Y_1 \rangle}$

Note that the rules that license the overt realization of valents do not alter the SLASH value. For now, we can simply assume that in these structures the mother node is required to have the same SLASH value as the head daughter. Thus, (58) behaves like an intransitive verb, and the information about the gap is 'percolated' as in Figure 1.

There is independent evidence for gaps being recorded locally in this way, Bouma et al. (2001), Levine and Hukari (2006) and others argue (although the cyclic movement is also consistent with it). Languages such as Chamorro (Chung, 1998), Ewe (Collins, 1994), Icelandic (Maling and Zaenen, 1978), Kikuyu (Clements, 1984), French (Kayne and Pollock, 1978), Yiddish (Diesing, 1990), and many others exhibit specific phonological or morphosyntactic connectivity phenomena in the structures intervening between the overt extracted element and the extraction site. Consider for example the case of Irish, as reported in McCloskey (1979). The verb particle *goN* only occurs with verbal structures that do not contain gaps, while *aL* only occurs between a filler and a gap:[13]

(61) a. Shíl m goN mbeadh s ann.
thought I VPART would-be he there
'I thought that he would be there'

b. An fear aL shíl m aL bheadh _ ann.
[the man] VPART thought I VPART would-be there
'the man that I thought would be here'

[13]In some accounts these particles are taken to be complementizers. The N orthographically indicates nasal mutation and L indicates lenition. I ignore the fact that these correspond to various surface forms which inflect for tense and negation.

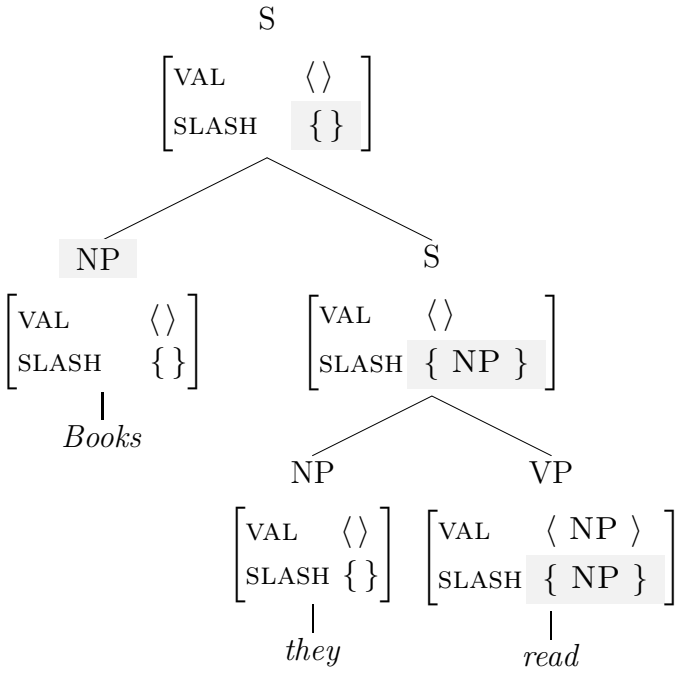

Figure 1: Representation of *Books, they read*

This evidence shows that lexical heads can detect the presence of unbounded dependencies in their dependents (Bouma et al., 2001). In the case of Irish, for example, one can assume that the verb particle *goN* selects [SLASH {}] clauses while *aL* selects [SLASH {...}] clauses (alternatively, the complementizer combines with the matrix verb and inspects the SLASH values of its valents).[14]

Let us turn to coordination, and to Gazdar's account of the CSC. Coordination is modeled along the lines of Yngve (1960) and others, with a coordination rule X $\rightarrow$ X_{crd-} X_{crd+} and a conjunct marking rule X_{crd+} $\rightarrow$ *crd-mrk* X_{crd-}. The feature *crd* identifies whether a conjunct is marked by a coordinator or not. These two rules license structures like the one in Figure 2.

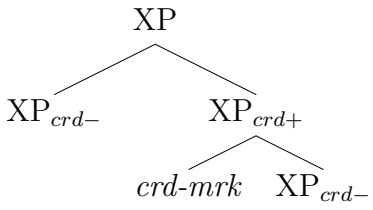

Figure 2: A generic coordinate structure

In Gazdar's account only identical categories can be conjoined, and thus, one cannot coordinate a transitive verb with a VP, or with a sentence. These have different valence

[14]See also Assmann et al. (2010). Note that we also have the necessary tools to deal with so-called *wh*-agreement languages like Chamorro and Palauan, if indeed such phenomenon even exists (Donohue and Maclachlan, 1999). Verbs receive a marker *-um-* when subjects are extracted, remain unchanged or nominalize if an object is extracted (if transitive, *-in* is added), if an oblique is extracted they nominalize (and if unaccusative, optionally use *-in-*). A cluster of rules for verbs can state the above constraints, and require that the verbal head agrees with a valent that is or contains a gap.

specifications and are therefore incompatible, as seen in (62). So-called coordination 'unlike categories' (Sag et al., 1985) can be modeled via underspecification (Pollard and Sag, 1994; Sag, 2002) or via ellipsis (Beavers and Sag, 2004; Chaves and Sag, 2007; Chaves, 2008). Nothing in my account hinges on this choice of analysis.

(62) a. *Fred [[read a book]$_{\text{VAL}\langle \text{NP} \rangle}$ [and [opened]$_{\text{VAL}\langle \text{NP, NP} \rangle}$]]

b. *Fred [[she has a hat]$_{\text{VAL}\langle \rangle}$ [and [smiled]$_{\text{VAL}\langle \text{NP} \rangle}$]]

Moreover, if one assumes X $\rightarrow$ X$_{crd-}$ X$_{crd+}$ requires that the syntactic features associated to the mother and daughter nodes must unify, then these valence lists are fused into one and the same list. In VP conjunction, for example, mother and daughters will contain the same VAL value, as Figure 3 shows.

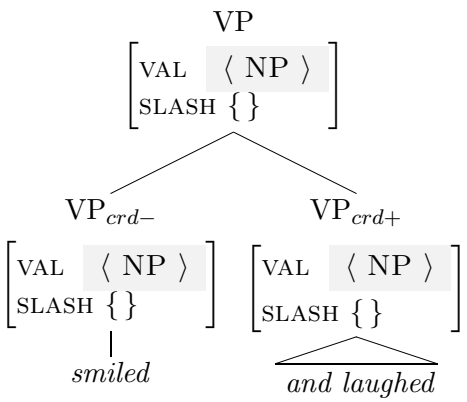

Figure 3: Argument sharing in coordination

Similarly, if coordination requires conjuncts to be identical with regard to SLASH values – as in Gazdar's account – then it follows that either all conjuncts contain the same gaps or no gaps at all. If conjuncts have different SLASH values then coordination fails as in (63b,c) and (64b,c).

(63) a. [To him]$_{\text{PP}}$, [Fred gave a football _]$_{\text{SLASH}\{\text{PP}\}}$ and [Kim gave a book _]$_{\text{SLASH}\{\text{PP}\}}$

b.*[To him]$_{\text{PP}}$, [Fred gave a football _]$_{\text{SLASH}\{\text{PP}\}}$ and [Kim gave me a book]$_{\text{SLASH}\{\}}$

c.*[To him]$_{\text{PP}}$, [Fred gave a football to me]$_{\text{SLASH}\{\}}$ and [Kim gave a book _]$_{\text{SLASH}\{\text{PP}\}}$

(64) a. It offers [something]$_{\text{NP}}$ [that every kid wants _]$_{\text{SLASH}\{\text{NP}\}}$ and [that every parent tries to help their child to achieve _]$_{\text{SLASH}\{\text{NP}\}}$

b.*It offers [something]$_{\text{NP}}$ [that every kid wants _]$_{\text{SLASH}\{\text{NP}\}}$ and [that every parent tries to help their child to achieve it]$_{\text{SLASH}\{\}}$

c.*It offers [something]$_{\text{NP}}$ [that every kid wants it]$_{\text{SLASH}\{\}}$ and [that every parent tries to help their child to achieve _]$_{\text{SLASH}\{\text{NP}\}}$

ATB extraction obtains when each conjunct contains a gap. The coordination rule yields a mother node in which all the daughter gaps are unified, just like local valents were unified in Figure 3. The topicalization example in Figure 4 illustrates gap unification in ATB extraction.

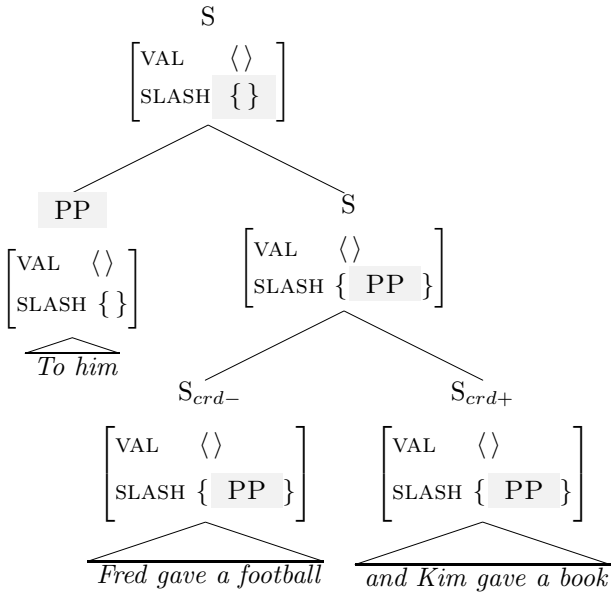

Figure 4: ATB extraction

This kind of analysis is superior to standard movement accounts because even though it resorts to a very simple mechanism (feature value equality), it is *simultaneously* compatible with constructions where there is one filler for multiple gaps, and constructions where there are multiple fillers for multiple gaps. Note that equality of feature values (also known as *structure-sharing*) is responsible for virtually all that goes on in an HPSG grammar, including subcategorization, semantic composition, agreement, morphology, phonology, pragmatics, etc.. It is a simpler and more pervasive concept than movement.

In spite of its relative superiority to transformational accounts, Gazdar's analysis of extraction in coordination structures cannot be right. Given that there is no evidence that asymmetric coordination is anything other than a kind of coordination structure, the question remains as to why ATB extraction is mandatory in symmetric coordination but optional in all other structures that we have discussed, headed or otherwise. The identity of SLASH values must not be triggered by the coordination construction, but by something else. In what follows I will discuss alternative explanations for why SLASH identity is only required by symmetric coordination.

4.3 Coherence and parallelism

Kuno (1987,23) argued that for an element to be extracted it must qualify as the topic of the sentence (see also Erteschik-Shir (1981) and Kehler (2002) for more discussion and related proposals). That is, it must be relevant for what the sentence is about, in some way. Consider (65), due to Kuno (1987). Sentence (65a) is about writing a book, books have topics, and therefore the writing action is directly connected to the book topic. In terms of Pustejovsky (1995,ch.5), for example, the book topic is a 'shadow argument' in the argument structure of *write*. Sentence (65b), however, is about losing a book, an action which does not directly extend to the book topic.

(65) a. Who did you write a book about? >

b. Who did you lose a book about?

Kuno's 'speaking of X' test identifies a potential sentence topic:

(66) a. Speaking of John Irving, I just bought a book about him. >

b. Speaking of John Irving, I just lost a book about him.

Based on these and other insights, Kehler (2002) proposes that symmetric coordination is associated to a pragmatic parallelism inference, which must hold even when elements are extracted. For an element to be placed in a topical position it must be able to serve as topic in each conjunct. Thus, extraction in symmetric coordination must be ATB because parallelism must be preserved, while extraction in asymmetric coordination need not be ATB because the underlying inferential pragmatic relation does not establish parallelism. In this view, the grammar of coordination would not have anything to say about extraction, and the ATB patterns become an epiphenomenon caused by pragmatic coherence conditions. Unfortunately, Kehler (2002) is not very explicit about how these inferences are construed.

Rather than resorting to pragmatic inferences, Goldsmith (1985) argued that extraction is parallel only when the *semantics* is also parallel. I agree with this position and hold that a proper account of the parallelism inference resides at the interface between the semantics of conjunction and Kuno's topicality observation. In what follows I will expand on this idea and propose a more explicit pragmatic-semantic explanation for why such parallelism exists in symmetric coordination. I start by questioning the assumption that the differences between symmetric and asymmetric coordination are due to pragmatics, and then show how a semantic account could deal with the facts.

Consider the asymmetric temporal-precedence *and*. It is standardly assumed that this temporal-precedence import is pragmatic in nature because it also appears when the same clauses are in parataxis. However, temporal precedence is truth-conditional, as first pointed out by Cohen (1971, 58). For example, (67a) and (67b) are not equivalent. The truth-conditions of the consequent clause crucially depend on the order of the conjuncts in the antecedent.

(67) a. If the old king has died of a heart attack and a republic has been declared, then Tom will be quite content.

b. If a republic has been declared and the old king has died of a heart attack, then Tom will be quite content.

Similarly, the examples in (68) taken from Wilson (1975, 151) and Levinson (2000, 206) suggest that this kind of asymmetric conjunction meaning is also truth-conditionally available to comparatives. I add to this evidence the example in (69), where *the other way around* is only felicitous if the conjunction corresponds to temporal precedence.

(68) a. Getting married and having kids is better then having kids and getting married.

b. If you have a baby and get married, then the baby is strictly speaking illegitimate.

(69) Should I soak it and scrub it, or the other way around?
(= should I scrub it and then soak it?)

Moreover, it is also the case that some asymmetric meanings do not show up in parataxis. For instance, the conditional interpretation in (70) is lost, the second sentence in (71) is stronger than its coordinate counterpart, the non-coordinate sentence in (72) is not coherent, and the semantics of the examples in (73) is radically different.

(70) You drink one more beer and you'll be too drunk to drive. $\neq$
You drink one more beer. You'll be too drunk to drive.

(71) I can drink two bottles of wine and not get drunk. $\neq$
I can drink two bottles of wine. I can't get drunk.

(72) Even Robin failed, and he is the smartest student in our class. $\neq$
Even Robin failed. He is the smartest student in our class.

(73) You can't get a new car and Kim get just a postcard. $\neq$
You can't get a new car. Kim can't get just a postcard.

I do not deny the existence of paratactic pragmatic inferences, but the data suggest that there is some semantic contribution in asymmetric conjunction. This is not surprising, since we can identify many other different senses for *and* and *or*, often accompanied by peculiar syntactic constraints on the conjuncts. For example, *and* can be interpreted as plurality-formation (74a), *X-er and X-er* intensification (74b), arithmetical addition (74c), numeral formation (74d), and packaging (74e). I view all of these conjunctions, including asymmetric conjunctions like temporal precedence or causal conjunction, as semantically distinct forms of conjunction.

(74) a. [A man and a woman] who married illegally were both arrested.
b. The sound became [louder and louder].
c. Let us suppose that [two and two] is five.
d. There were [two hundred and one] UFO sightings this year.
e. [Two ham rolls and a glass of milk] was more than Sue wanted.

In what follows I show how the parallel/ATB extraction patterns can be obtained as a prediction of Kuno's topicality condition and the semantics of symmetric coordination. I start by proposing a pragmatic predicate $R(x, e)$ that requires that a filler x must be pragmatically relevant/coherent in the event described by the sister phrase e. The intended definition of R should be intuitive from the discussion in Kuno (1987). However, I propose that R is transitive, as defined in (75).

(75) **Filler-Head Relevance Condition:**
$R(x, e)$ holds *iff* i) x is a coherent argument of event e; or ii) there is some $P(e, ..., e', ...)$ such that $R(x, e')$ holds.

For example, in a sentence like *who*$_x$ *left*$_e$*?* the constraint $R(x, e)$ can be satisfied if the argument x is relevant for the verb that describes e. People frequently leave, and so there is nothing incoherent about this example. Conversely, in *who*$_x$ *do you think*$_e$ *left*$_{e'}$*?* the condition $R(x, e)$ cannot be trivially satisfied because x is not an argument of *think*. However, this verb takes e' as an argument, and therefore the condition $R(x, e')$ can hold.

Let us now turn to the semantics of coordination for a moment. The symmetric plurality-formation conjunction illustrated in (74a) obtains a plural NP $[[a\ man]_{x_1}$ *and* $[a\ woman]_{x_2}]_{x_1+x_2}$. Here, I use the operator '+' to denote to a mereological sum (Link, 1984). Predicates like *smile* or *happy* apply distributively to each member of a plurality, whereas predicates like *marry*, *flock*, *hire*, and *between* can apply to a plurality as a whole. Unlike the other conjunctions in (74), plurality-formation conjunction can combine daughters of any category. In particular, it can form *event pluralities* from verbal conjuncts (Bach, 1986; Eckardt, 1995; Lasersohn, 1995; Link, 1998). The sentence in (76a) – adapted from Oehrle (1987) – describes the frequency of two joint event-types, not of independent frequencies of occurrence. Adverbs like *simultaneously* and *alternately* apply non-distributively to event pluralities, while adverbs like *often* allow distributive and non-distributive readings.

(76) a. Often $[[$I go to the beach$]_{e_1}$ and [you go to the city$]_{e_2}]_{e_1+e_2}$.

b. I can't simultaneously $[[$drive a car $]_{e_1}$and [talk on the phone$]_{e_2}]_{e_1+e_2}$.

c. Sue $[[[$read the instructions$]_{e_1}$ and [dried her hair$]_{e_2}]_{e_1+e_2}$, [in exactly twenty seconds]].

d. He spent the day [alternately [criticizing Sue$]_{e_1}$ and [being criticized by her$]_{e_2}]_{e_1+e_2}$.

Thus, nominal conjuncts can form a nominal plurality and verbal conjuncts can form an event plurality. When a predicate applies to a plurality it can either predicate it directly (a non-distributive reading) or extend distributively to its mereological parts (distributive reading). More formally, a predicate P applying to some argument x can either apply directly as in $P(x)$, or apply to its mereological parts $P(x_1) \wedge \ldots \wedge P(x_n)$ if $x = x_1 + \ldots + x_n$. Similarly for symmetric disjunction: $P(x_1) \vee \ldots \vee P(x_n)$ is equivalent to $P(x)$ if $x = x_1 \vee \ldots \vee x = x_n$. I propose that essentially the same distributivity mechanism is at work in R, in the sense that it distributes to the conjuncts/disjuncts that it applies to. Hence, when R applies to an event sum $R(x, e_1 + e_2)$ then it must be the case that $R(x, e_1) \wedge R(x, e_2)$ holds. This semantic-pragmatic distribution predicts ATB extraction: the relevance condition is distributed over each conjunct, just like other predications apply distributively to pluralities. Asymmetric coordination meanings cannot force distributivity, and consequently, R does not have to distribute, thus explaining why ATB extraction is not obligatory in asymmetric coordination (see §5.3 for more discussion).

5 On the Syntax of Extraction

Head-Driven Phrase Structure Grammar (Pollard and Sag, 1994) is an ideal framework for stating a theory of extraction consistent with the empirical observations

made so far. First, being a declarative and non-derivational theory, HPSG is process-independent. The grammar is composed of a set of conditions – some language-specific, others universal – that characterize the set of grammatical expressions. The grammar does not generate or derive any structure. Rather, the grammar consists of a set of declarative axioms which characterize the admissible linguistic entities. Actual sentence production and understanding is assumed to be carried out by a psycholinguistically motivated language processing module – which is conditioned by cognitive limitations, such as finite memory resources – and uses the grammar rules in order to parse linguistic expressions (Sag and Wasow, 2011). This means that HPSG can in principle be coupled with theories of human sentence processing that model the kinds of performance limitations and biases discussed so far (see for example Konieczny (1996)).

A second reason for adopting a theory like HPSG is that the feature equality mechanism (i.e. structure-sharing) that is used model the gap propagation and linkage patterns discussed so far is also behind virtually everything all that goes on in the grammar, including agreement, case assignment, scope, variable binding, phonology, etc.. Finally, being a lexicalist and construction-based framework, HPSG enables constructions and lexical expressions to impose construction-specific constraints on extraction while maintaining that the syntactic mechanism that handles extraction in general operates in a uniform way in all constructions. Thus, we can capture the fact that comparatives and *tough*-movement adjectives, impose special extraction patterns. In what follows I put forth a general account that reconciles the extraction phenomena observed in symmetric and non-symmetric coordination with the extraction phenomena seen in non-coordinate structures. Due to space limitations I cannot discuss in detail how semantic representations are obtained in HPSG, and refer the reader to Copestake et al. (2006).

Ginzburg and Sag (2000) and Bouma et al. (2001) address various empirical problems with the theory of extraction in Pollard and Sag (1994, ch. 4) and propose an account where there are two kinds of signs – 'gap signs' and 'non-gap signs'. The interaction of several grammar rules determines how unbounded dependencies are formed: i) one rule states that gap-signs introduce a gap, ii) a 'canonicality principle' requires that only non-gap signs can be realized *in situ*, iii) a 'SLASH amalgamation' rule states that the gaps of a lexical head are by default the optional fusion of the gaps of the dependents, and iv) the ARGUMENT REALIZATION PRINCIPLE allows any complement to be a gap. The cancellation of gaps from SLASH can be done by phrasal constructions (e.g. via the topicalization rule) or lexically (e.g. the lexical entry of *tough*-movement words).

However, there are problems when extending this grammar to extraction in non-headed constructions, because the constraints responsible for gap percolation/fusion are stated in terms of *heads*: the SLASH value of a word contains the (optionally fused) gaps of its dependents, and gap percolation is achieved by stating that the mother node and the head daughter have the same SLASH values. Since gap fusion is handled at the lexical level, the theory fails to account for fusions that occur between gaps located in non-headed structures, such as non-symmetric coordination and comparatives. A more general approach is needed precisely because these constructions allow the same extraction patterns that subordinate structures allow (see §2.1). I propose a simpler

alternative account of the extraction mechanism where a single general phrasal rule is responsible for the percolation, optional fusion, and discharge of gaps in headed and non-headed constructions alike.

It should be noted that the problem of extraction in non-headed structures is also present in earlier versions of HPSG, including Pollard and Sag (1994, 164). There, the rule that governs gap percolation/fusion states that the mother's SLASH value is defined as the union of the daughters's SLASH values minus the the TO-BIND value of the *head daughter*. In order to achieve a uniform theory of extraction that aims to encompass both headed and non-headed constructions, I reject the feature TO-BIND, and abandon several of the assumptions in Ginzburg and Sag (2000) and Bouma et al. (2001).

I begin by replacing the notion of 'gap signs' and 'non-gap signs' with the notion of percolating gaps (SLASH signs which have not been linked to a filler) and filled gaps (SLASH signs which have been linked to a filler, and thus cease to percolate). I will also reformulate the ARGUMENT REALIZATION PRINCIPLE so that the extraction of any valents is permitted. Finally, the SLASH value of any phrase will be defined as the union of the percolating gaps located in SLASH values of the daughters. In my account, I return to Pollard and Sag (1994, 164) in assuming that gap propagation does not involve heads until the point in the structure when an element in SLASH is cashed out.

5.1 Basics

In what follows I adopt in general terms the feature geometry of Sign-Based Construction Grammar (Sag, 2010b), a constructional approach within the HPSG framework. In this theory, a distinction is made between signs (complex linguistic entities that simultaneously contain phonological, syntactic, and semantic information) and constructions (the rules that license the combination of signs in local tree structures). Both signs and constructions are encoded as typed feature structure descriptions. Take for example the verb *likes* in (77).

(77)
$$\begin{bmatrix} \textit{word} \\ \text{PHON} \ \langle\ \textit{likes}\ \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \textit{verb} \\ \text{VFORM}\ \textit{fin} \end{bmatrix} \\ \text{VAL}\ \ldots \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{INDEX} \quad \boxed{e} \\ \text{FRAMES}\ \langle\, \textit{like}(\boxed{e}, \boxed{x}, \boxed{y})\, \rangle \end{bmatrix} \\ \text{SLASH}\ \ldots \\ \text{ARG-ST}\ \langle \text{NP}[\text{INDEX}\ \boxed{x}],\ \text{NP}[\text{INDEX}\ \boxed{y}] \rangle \end{bmatrix}$$

The feature PHON contains phonology (simplified here for exposition), the feature SYN contains syntactic information about category, the verb form, and valence, and SEM contains semantics. A feature INDEX singles out the main index of a sign and a feature FRAMES which lists the predications that jointly describe the meaning of a sign. For exposition purposes I depict semantic predications as first-order terms rather than as features.

The ARG(UMENT)-ST(RUCTURE) feature contains the list of arguments that the word selects, ordered by obliqueness. This list is relevant for Binding Theory, among other things. The verb in (77) selects a subject NP and a complement NP, and the respective indices associated to these arguments are co-indexed with the respective argument slots in the semantics: $like(e, x, y)$. The notation NP[INDEX $\boxed{x}$] is used to abbreviate a nominal structure with an empty VAL(ENCE) list and a semantic index x. Notice that the VAL(ENCE) and SLASH features are underspecified in these lexical entries. A rule called ARGUMENT REALIZATION PRINCIPLE will determine the various ways in which the elements in ARG-ST can be realized in VAL and SLASH.

Signs and constructions are hierarchically organized in a system of types. Thus, any constraint imposed over a linguistic structure of a given type τ must also hold for all the subtypes of τ. Figure 5 shows a (simplified) type hierarchy for the type *sign*, with the types for (un)inflected lexemes omitted.

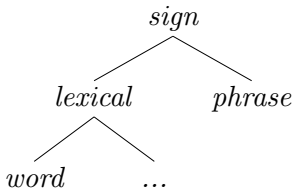

Figure 5: Type hierarchy of signs

Lexical and phrasal constructions are represented with trees, or more precisely, with feature structures that encode trees as in Figure 6. The feature MTR corresponds to the mother node of the structure, the list DTRS contains the local daughters, and HD-DTR singles out the head daughter. The trees that the grammar licenses are called constructs. The occurrence of boxed tags like $\boxed{1}$ signifies that the two feature structures are identical. Feature value equality is called *structure-sharing*, and is responsible for propagating all kinds of information in the grammar, including subcategorization, binding, linearization, semantic composition, phonology, morphology, and extraction.

$$\begin{array}{c} \text{X} \\ \text{Y} \quad \text{Z} \end{array} = \begin{bmatrix} \textit{construct} & \\ \text{MTR} & \begin{bmatrix} \textit{sign} \\ \text{SYN X} \end{bmatrix} \\ \text{HD-DTR} & \boxed{1} \\ \text{DTRS} & \left\langle \begin{bmatrix} \textit{sign} \\ \text{SYN Y} \end{bmatrix}, \boxed{1} \begin{bmatrix} \textit{sign} \\ \text{SYN Z} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 6: A feature structure representation of a tree headed by the second daughter

Only structures of type *word* or *phrase* are allowed in phrasal constructs, as made explicit in rule (78). The members of DTRS must be words or phrases.

(78) $$\textit{phrasal-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \textit{phrase} \\ \text{DTRS} & \textit{list}(\textit{word} \vee \textit{phrase}) \end{bmatrix}$$

The inventory of phrasal types relevant for this work is in Figure 7, all of which must obey (78). Each subtype is motivated by particular syntactic and/or semantic properties peculiar to that kind of construction. There are five major constructs: head-complement constructs (*pred-hd-comp-cxt*; for verbal heads and their complements), head-subject constructs (*subj-pred-cxt*; for verbal heads and their subjects), head-filler constructs (*filler-hd-cxt*; for a filler and a gapped clause), head-functor constructs (*hd-func-cxt*; for modifiers/markers and their heads), and coordinate constructs (*coord-cxt*).

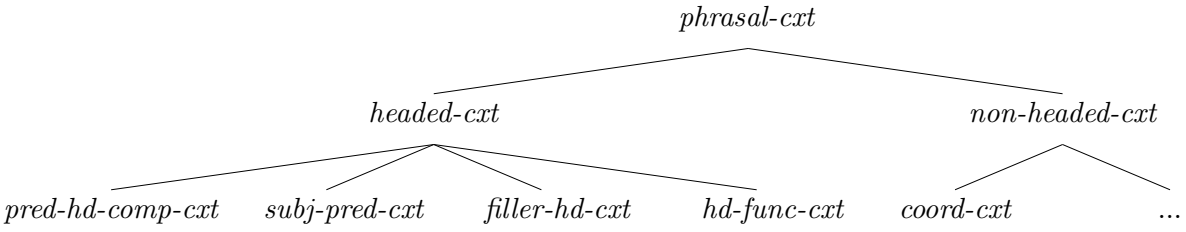

Figure 7: Type hierarchy of lexical and phrasal constructions

The *pred-hd-comp-cxt* rule is provided in (79) below, for illustration. This rule states that any tree structure of the type *pred-hd-comp-cxt* is composed of a lexical verbal daughter (2) and its complements (3). The only valent that the mother and the head share is the subject 1. More specifically, the list of daughters of this construction is split into two non-empty sub-lists (via the list concatenation relation '$\oplus$'), one containing the verbal head 2 and another (3) containing all the complements required by the head. Thus, the members of the list 3 in VAL are required to be members of DTRS, as sisters of the head.[15]

(79) PREDICATION HEAD-COMPLEMENT CONSTRUCTION

$$\textit{pred-hd-comp-cxt} \Rightarrow \left[\begin{array}{l} \textsc{mtr}\left[\textsc{syn}\left[\textsc{val}\ \langle \boxed{1} \rangle\right]\right] \\ \textsc{hd-dtr}\ \boxed{2} \\ \textsc{dtrs}\left\langle \boxed{2}\left[\begin{array}{ll} \textit{word} & \\ \textsc{syn} & \left[\begin{array}{l}\textsc{cat}\ \textit{verb} \\ \textsc{val}\ \langle \boxed{1} \rangle \oplus \boxed{3}\end{array}\right]\end{array}\right]\right\rangle \oplus \boxed{3}_{\textit{nelist}} \end{array}\right]$$

Although nothing in my account hinges on this move, a different phrasal rule is assumed to deal with complementation structures headed by elements that do not select specifiers (such as prepositions and relational adjectives like *proud*). For more details about this feature geometry see Sag (2010b).

Another example of a rule that captures a generalization about a set of constructions is the HEAD FEATURE PRINCIPLE, which ensures that in headed constructions the part-of-speech of the mother node is the same as the head daughter's. More precisely, the same CAT value:

[15]List concatenation can be expressed as simple structure-sharing (Copestake et al., 2001).

(80) Head Feature Principle

$$\textit{headed-cxt} \Rightarrow \begin{bmatrix} \text{MTR} \mid \text{SYN} \mid \text{CAT} & \boxed{1} \\ \text{HD-DTR} \mid \text{SYN} \mid \text{CAT} & \boxed{1} \end{bmatrix}$$

The rules for the other constructions in Figure 7 are provided below, as the grammar fragment for extraction is fleshed out.

5.2 Extraction

I begin by assuming that the members of SLASH are not signs, but entities of the type *gap*. This type introduces a feature GAP which records the extracted sign, as defined in (81). There are also two types of gap: *pg* (percolating gap) and *fg* (non-percolating filled gap). Only the former type will be allowed to percolate in the tree structure.

(81)

$$\begin{bmatrix} \text{SLASH} & \textit{set}\left(\begin{bmatrix} \textit{gap} \\ \text{GAP} \;\; \textit{sign} \end{bmatrix}\right) \end{bmatrix}$$

In order to allow any valents to be extracted, I propose the Argument Realization Principle (ARP) in (82).[16] This rule states that the elements in ARG-ST are the same as the ones in VAL, but optionally, valents can be removed from VAL, and placed in SLASH as percolating gaps (*pg*). The valency reduction operation is implemented via the list subtraction operation '$\ominus$' (Ginzburg and Sag (2000,170,ft.7) which allows any sublist of n ($0 \leq n$) elements to be removed from the argument structure list $\boxed{1}$. Nothing requires that the subtracted elements X_0 ... X_n occur adjacently in $\boxed{1}$.

(82) Argument Realization Principle:

$$\textit{word} \Rightarrow \begin{bmatrix} \text{SYN} & \left[\text{VAL} \;\; \boxed{1} \ominus \langle \boxed{X_0}, \ldots, \boxed{X_n} \rangle\right] \\ \text{SLASH} & \left\{ \begin{bmatrix} \textit{pg} \\ \text{GAP} \;\; \boxed{X_0} \end{bmatrix}, \ldots, \begin{bmatrix} \textit{pg} \\ \text{GAP} \;\; \boxed{X_n} \end{bmatrix} \right\} \\ \text{ARG-ST} & \boxed{1} \end{bmatrix}$$

The ARP can take as input the verb (77) and produce, for example, (83a) or (83b). The subcategorized items must be realized *in situ* in the former, while the complement is extracted in the latter. In English, the valents of any word can undergo this process because VAL and SLASH are lexically underspecified.[17]

[16] Cf. with Ginzburg and Sag (2000, ch.5) and Levine and Hukari (2006, 104).

[17] In languages without preposition stranding, for example, the SLASH values of prepositions are specified as [SLASH {}], and thus the ARP cannot allow any valent to reside in SLASH.

(83) a.
$$\begin{bmatrix} \textit{word} \\ \text{PHON} \ \langle\ \textit{likes}\ \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \textit{verb} \\ \text{VFORM}\ \textit{fin} \end{bmatrix} \\ \text{VAL}\ \langle \boxed{1}, \boxed{2} \rangle \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{INDEX} \quad \boxed{e} \\ \text{FRAMES}\ \langle\ \textit{like}(\boxed{e}, \boxed{x}, \boxed{y})\ \rangle \end{bmatrix} \\ \text{SLASH}\ \{\} \\ \text{ARG-ST}\ \left\langle \boxed{1}\text{NP}[\text{INDEX}\ \boxed{x}],\ \boxed{2}\text{NP}[\text{INDEX}\ \boxed{y}] \right\rangle \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{word} \\ \text{PHON} \ \langle\ \textit{likes}\ \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \textit{verb} \\ \text{VFORM}\ \textit{fin} \end{bmatrix} \\ \text{VAL} \quad \langle \boxed{1} \rangle \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{INDEX} \quad \boxed{e} \\ \text{FRAMES}\ \langle\ \textit{like}(\boxed{e}, \boxed{x}, \boxed{y})\ \rangle \end{bmatrix} \\ \text{SLASH} \left\{ \begin{bmatrix} \textit{pg} \\ \text{GAP}\ \boxed{2} \end{bmatrix} \right\} \\ \text{ARG-ST}\ \left\langle \boxed{1}\text{NP}[\text{INDEX}\ \boxed{x}],\ \boxed{2}\text{NP}[\text{INDEX}\ \boxed{y}] \right\rangle \end{bmatrix}$$

I also assume that most words in the lexicon must be of the form shown in (84). The exceptions are those that trigger *tough*-movement, to be discussed shortly. This constraint simply ensures that all gaps located in phrasal arguments also start out as percolating gaps, and is necessitated by how the percolation and discharge of SLASH values is handled.

(84)
$$\left[\text{ARG-ST}\ \textit{list}\left(\left[\text{SLASH}\ \textit{set}(\textit{pg}) \right] \right) \right]$$

The percolation, fusion, and discharge of any number of gaps, in any syntactic construction, with any number of k daughters, is modeled by the GENERALIZED NON-LOCAL INHERITANCE PRINCIPLE (GNIP) in (85). This rule states that in every construction, any daughter may contain a set of any number of percolating gaps (i.e. typed *pg*) and a set of any number of filled gaps (typed *fg*). Only percolating gaps percolate to the mother, and possibly fuse. Note that all of the gaps in the mother node are type-underspecified as *gap*, thus allowing subsequent structures to type them as *fg* or as *pg*.

(85) Generalized Non-Local Inheritance Principle

$$\textit{phrasal-cxt} \Rightarrow \left[\begin{array}{l} \text{MTR}\left[\text{SLASH}\left\{\left[\begin{array}{l}gap\\ \text{GAP}\ \boxed{X_0^0}\end{array}\right],\ldots,\left[\begin{array}{l}gap\\ \text{GAP}\ \boxed{X_0^n}\end{array}\right]\right\}\cup\ldots\cup\left\{\left[\begin{array}{l}gap\\ \text{GAP}\ \boxed{X_k^0}\end{array}\right],\ldots,\left[\begin{array}{l}gap\\ \text{GAP}\ \boxed{X_k^m}\end{array}\right]\right\}\right] \\ \text{DTRS}\left\langle \begin{array}{c}\left[\text{SLASH}\left\{\left[\begin{array}{l}pg\\ \text{GAP}\ \boxed{X_0^0}\end{array}\right],\ldots,\left[\begin{array}{l}pg\\ \text{GAP}\ \boxed{X_0^n}\end{array}\right]\right\}\cup set(fg)\right] \\ \vdots \\ \left[\text{SLASH}\left\{\left[\begin{array}{l}pg\\ \text{GAP}\ \boxed{X_k^0}\end{array}\right],\ldots,\left[\begin{array}{l}pg\\ \text{GAP}\ \boxed{X_k^m}\end{array}\right]\right\}\cup set(fg)\right]\end{array}\right\rangle \end{array}\right]$$

Let us consider a few examples of how the GNIP achieves gap percolation and fusion. Since the *pred-hd-comp-cxt* rule in (79) does not impose any constraints on SLASH values, then any daughter can contain gaps. This is shown in Figure 8, for the VP in sentence (86).

(86) This is the person who$_j$ I can't remember which papers$_i$ I [sent [copies of $\__i$] [to $\__j$]]$_{\text{VP}}$.

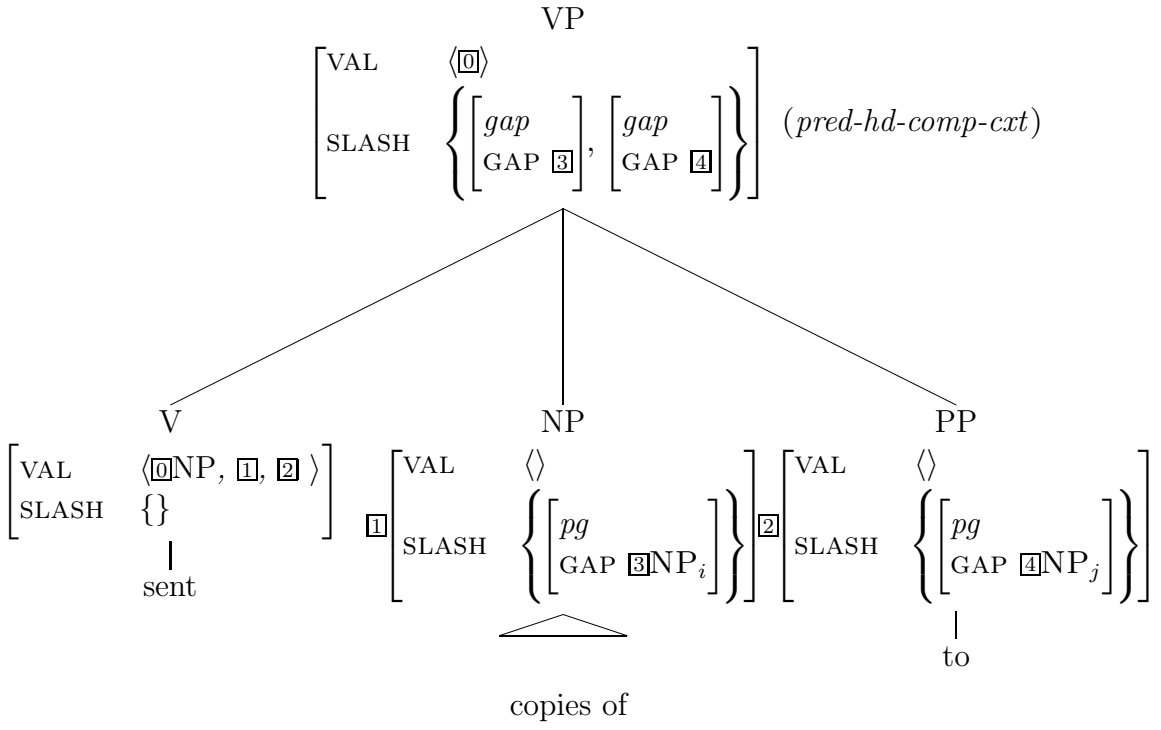

Figure 8: Multiple gap percolation

In Figure 8, the gaps are simply percolated, but in the VP in (87) the gaps are fused and then percolated. This is depicted in Figure 9. Both of these cases are due to the ARP allowing the prepositional complements to be percolating gaps rather than

in situ complements. When the signs in GAP are co-referential and correspond to syntactically identical categories, then set union ensures that they are one and the same entity.[18]

(87) I can't remember who$_i$ I [sent [nude photos of $\__i$] [to $\__i$]]$_{\text{VP}}$.

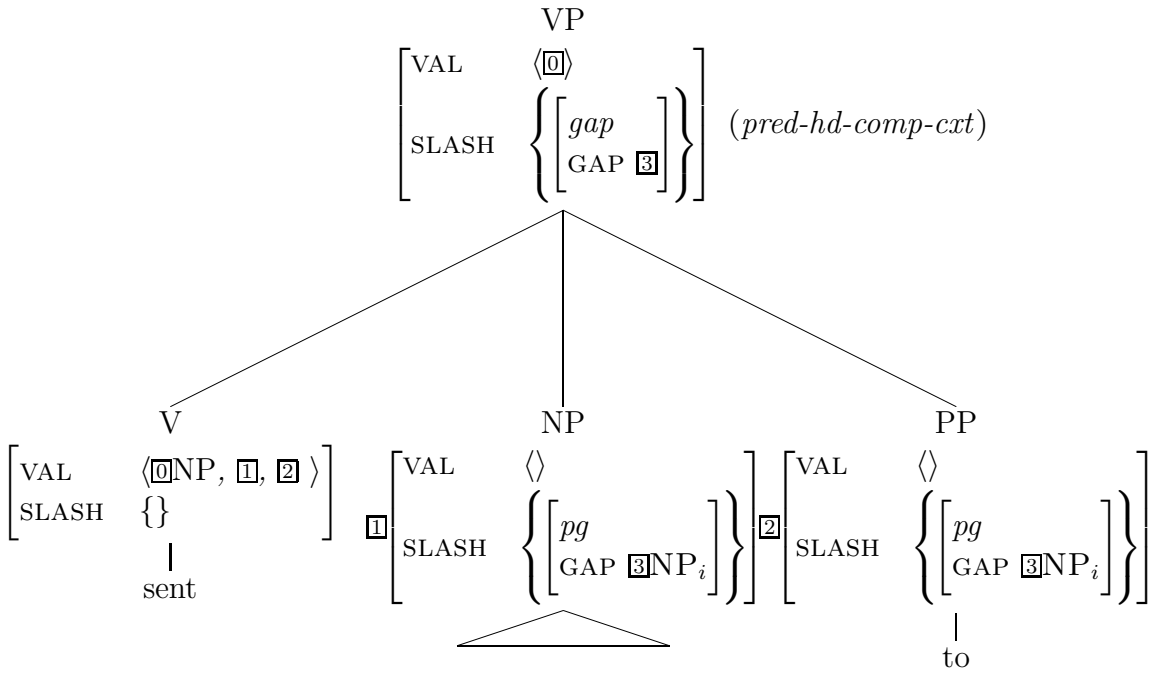

Figure 9: Gap fusion and percolation

Various other extraction possibilities are allowed by the ARP. It is possible that only the verb is gapped (*this is the person who$_i$ I [sent $\__i$ nude photos of you]*), that only the preposition is gapped (*this is the person who$_i$ I [sent you nude photos of $\__i$]*), or finally, that both the verb *sent* and the preposition *of* are gapped with independent gaps, as in *this is the person who$_i$ I can't remember what$_j$ I [sent $\__i$ [photos of $\__j$]*.

The rule that allows subjects to combine with verbal phrases is (88). Any structure of the type *subj-pred-cxt* consists of a verbal head daughter 2 that selects a subject 1. Any gaps recorded by the head daughter must be typed as percolating. Similarly, any gaps recorded in the subject daughter are also typed as percolating because of the constraints in (84).

[18]It seems that – in certain conditions – gap fusion can cumulate rather than identify gaps, as in *what$_{i+j}$ did you say that Tom drank $\__i$ and ate $\__j$?* This can be captured by redefining the set union operator to allow either identity or summation of set members, along the lines of Chaves (to appear). For more data, different perspectives and controversy see Postal (1998, 136,160), Munn (1998, 1999) and Gawron and Kehler (2003).

(88) Subject-Predicate Construction

$$\textit{subj-pred-cxt} \Rightarrow \left[\begin{array}{l} \textsc{mtr}\left[\textsc{syn}\left[\textsc{val}\ \langle\rangle\right]\right] \\ \textsc{hd-dtr}\ \boxed{2} \\ \textsc{dtrs}\left\langle \boxed{1}, \boxed{2}\left[\begin{array}{l} \textsc{syn}\left[\begin{array}{l}\textsc{cat}\ |\ \textsc{vform}\ \textit{fin} \\ \textsc{val}\ \langle\boxed{1}\rangle\end{array}\right] \\ \textsc{slash}\ \textit{set}(pg)\end{array}\right]\right\rangle \end{array}\right]$$

In order to see the grammar fragment at work, let us focus on topicalization structures and assume the rule in (89). See Sag (2010a) for a formalization of a larger class of filler-gap constructions, where topicalization constructions are included. The rule in (89) requires that the head daughter $\boxed{2}$ contains exactly one *fg* gap in SLASH, and that $\boxed{1}$ is the filler. The gap typing automatically ensures that the two subsets in SLASH are disjoint.

(89) Head-Filler Construction

$$\textit{filler-hd-cxt} \Rightarrow \left[\begin{array}{l} \textsc{mtr}\,|\,\textsc{syn}\,|\,\textsc{val}\ \langle\ \rangle \\ \textsc{hd-dtr}\ \boxed{2} \\ \textsc{dtrs}\left\langle \boxed{1}, \boxed{2}\left[\begin{array}{l} \textsc{syn}\left[\begin{array}{l}\textsc{cat}\ \textit{verb} \\ \textsc{val}\ \langle\ \rangle\end{array}\right] \\ \textsc{slash}\left\{\left[\begin{array}{l}\textit{fg} \\ \textsc{gap}\ \boxed{1}\end{array}\right]\right\} \cup \textit{set}(pg)\end{array}\right]\right\rangle \end{array}\right]$$

A topicalization example is provided in Figure 10. The gap starts out as *pg* because of the lexical entry in (83b). The GNIP percolates the gap in every node where phrasal rules require gaps to be typed as *pg*. The head-filler rule requires a gap typed as *fg* and the GNIP prevents its percolation.

Since the Head-Filler Construction in (89) does not impose any conditions on the filler's SLASH value, it allows extractions out of filler daughters as shown in (90). Following Sag (2010a), I assume extraction from topicalized constituents is blocked by the topicalization rule (a sub-type of *hd-filler-cxt*), which requires among other things that the filler daughter is [SLASH {}].

(90) This is [the handout]$_j$ that I can't remember [how many copies of _ $_j$]$_i$ we have to print _ $_i$.

With regard to so-called *tough*-movement, I assume that words like *easy* and *eager* select an infinitive complement VP that must contain one gap typed as *fg* (i.e. no percolation). This gap is referentially linked to the external argument of the adjective, and any other gap in the VP must be of the percolating kind, as shown in (91). Thus, **who*$_i$ *did you think Tom was easy* [*to talk to* _ $_i$] is out because the *fg* gap in the VP is unable to percolate beyond the AP phrase (since it is typed as *fg*), and **Tom is easy* [*to talk to Fred*] is out because that VP does not have a gap. The [SLASH

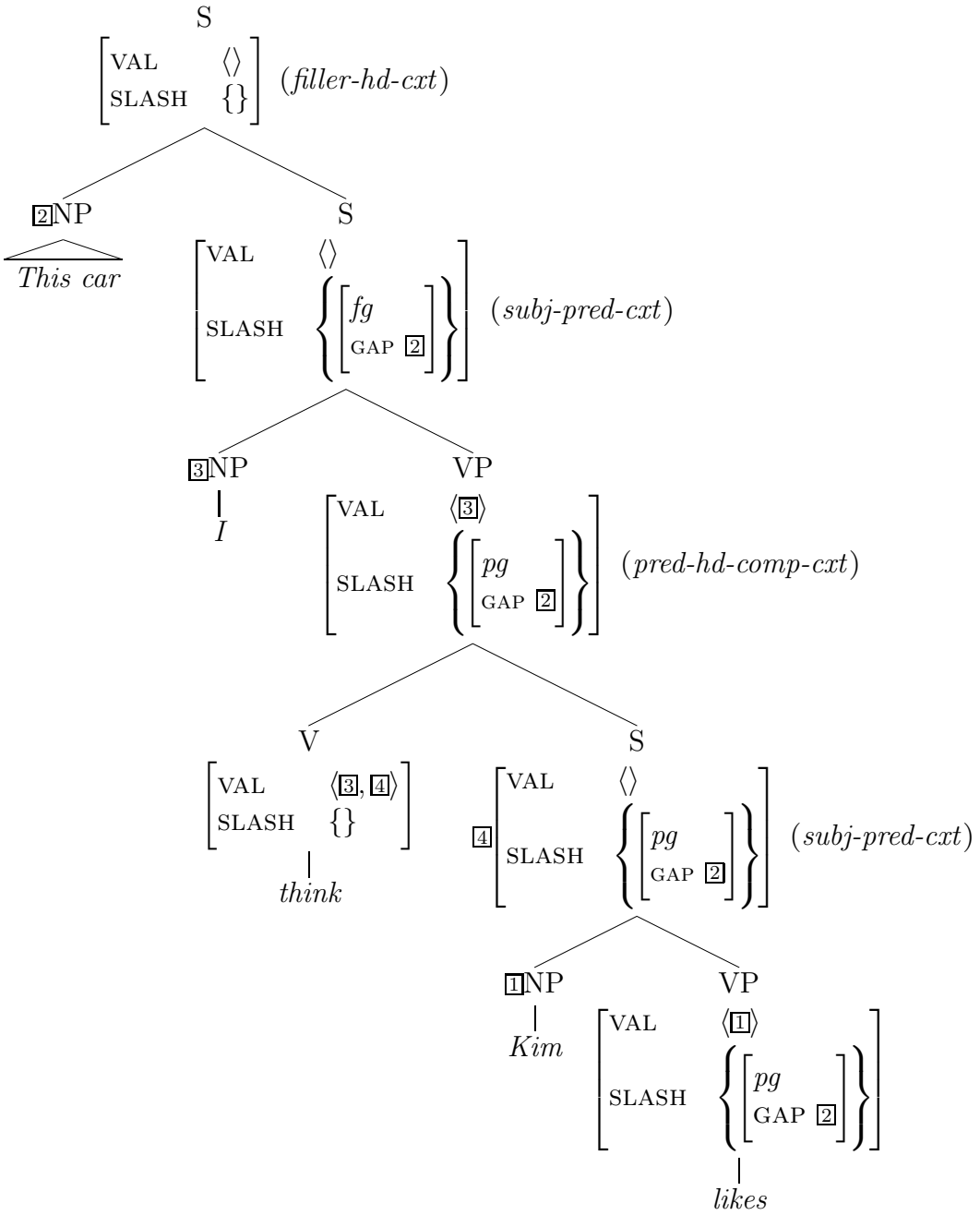

Figure 10: A topicalization filler-gap dependency

{}] constraint ensures that the PP is an island (**it was HIM that Kim was easy for* _ *to please*). Following the analysis of extraposition in Kim and Sag (2005), I assume that the counterpart of the adjective that occurs in *it*-extraposition constructions is required to obey (84) instead, therefore disallowing any *fg* gaps.

(91)
$$\begin{bmatrix} \textit{word} \\ \text{PHON} \ \langle \ \textit{easy} \ \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} \begin{bmatrix} \textit{adj} \\ \text{XARG} \ [\text{INDEX} \ \boxed{x}] \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} \left\langle (\text{PP}\left[\text{SLASH} \ \{\}\right],)\text{VP}_{\textit{inf}} \left[\text{SLASH} \left\{ \begin{bmatrix} \textit{fg} \\ \text{GAP}\left[\text{INDEX} \ \boxed{x}\right] \end{bmatrix} \right\} \cup \textit{set}(pg) \right] \right\rangle \end{bmatrix}$$

Similarly to *subj-pred-cxt* and *pred-hd-comp-cxt*, modification structures do not impose any constraints on SLASH. This will allow any daughter to contain gaps (fused or otherwise), since the parasitic effects are due to non-grammatical factors (§3). Following Van Eynde (2003), I assume that adjuncts – and certain other expressions such as determiners, complementizers, and coordinators – have a feature SELECT that allows them to impose constraints on the head phrase that they attach to. Thus, prepositions, adjectives and relative clauses select a nominal head, adverbs select a verbal head, and so on. The relevant rule for these structures is (92), which states that a phrase of type *hd-func-cxt* can be formed with a non-head daughter and a head daughter, as long as the latter is compatible with the former's SELECT value. I assume that the linearization of adjuncts is captured via linear order constraints like those in Kathol (2000). Note that the valence of the head daughter and the mother are the same: $\boxed{1}$. This ensures that adjoining a modifier to some head produces a mother node of the same category as the head daughter. Also, certain 'marking' information – needed for complementizers and coordinators – is percolated via MRKG and CRD, respectively, as discussed below.

(92) HEAD-FUNCTOR CONSTRUCTION:

$$\textit{hd-func-cxt} \Rightarrow \begin{bmatrix} \text{MTR} \begin{bmatrix} \text{SYN} \begin{bmatrix} \text{CAT} \mid \text{SELECT} \ \textit{none} \\ \text{VAL} \quad \boxed{1} \\ \text{MRKG} \ \boxed{3} \end{bmatrix} \\ \text{CRD} \ \boxed{4} \end{bmatrix} \\ \text{HD-DTR} \ \boxed{2} \\ \text{DTRS} \left\langle \begin{bmatrix} \text{SYN} \begin{bmatrix} \text{CAT} \mid \text{SELECT} \ \boxed{2} \\ \text{VAL} \ \langle \, \rangle \\ \text{MRKG} \ \boxed{3} \end{bmatrix} \\ \text{CRD} \ \boxed{4} \end{bmatrix}, \boxed{2} \begin{bmatrix} \text{SYN} \mid \text{VAL} \ \boxed{1} \\ \text{SLASH} \ \textit{set}(pg) \end{bmatrix} \right\rangle \end{bmatrix}$$

For lack of space, I cannot discuss adjunct extraction here. I refer the reader to Hukari and Levine (1995), Bouma et al. (2001), Levine (2003), Sag (2005), Levine and Hukari (2006), Sato and Tam (2008) and Chaves (2009).

5.3 Coordination

In this section I focus in monosyndetic coordination, but the constructional framework that I am adopting can be scaled to other coordination patterns. Syntactically, I take

inspiration from Beavers and Sag (2004). Coordinators are allowed to attach to a phrase and mark them as [CRD *marked*], and all words in the lexicon are specified as [CRD *unmarked*]. In order to distinguish between different semantic kinds of coordinators, I propose a feature [MODE *coord-mode*] and the type hierarchy in Figure 11 for the type *coord-mode*. This type distinguishes between plurality-forming conjunction ('+'), Boolean disjunction ('∨'), temporal precedence ('≺'), causal conjunction ('→'), etc..

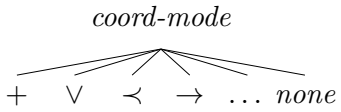

Figure 11: Semantic coordination types

The lexical entries for coordinators are given in (93). These have no valents because coordinators are markers, rather than heads. As a direct consequence, the ARP will not allow conjuncts to be extracted. Like adjuncts, determiners, and complementizers, coordinators select their host phrase via SELECT. The conjunction in (93a) corresponds to plurality formation via a mereological sum '+', the disjunction in (93b) it corresponds to classical disjunction '∨'. Both of these operators are dyadic, and their interpretation is briefly discussed below in (102). Note also that one need not assume a different lexical entry for each coordinator. The various meanings of *and* and *or* can easily be grouped in the same (type-underspecified) lexical entry by partitioning *coord-mode* types.

(93) a.
$$\begin{bmatrix} \textit{word} \\ \text{PHON} \ \langle \ \textit{and} \ \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} & \textit{coord} \\ \text{SELECT} & \begin{bmatrix} \text{CRD} \ \textit{unmarked} \end{bmatrix} \\ \text{VAL} & \langle\rangle \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{FRAMES} \ \langle\rangle \end{bmatrix} \\ \text{CRD} \begin{bmatrix} \textit{marked} \\ \text{MODE} \quad + \end{bmatrix} \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{word} \\ \text{PHON} \ \langle \ \textit{or} \ \rangle \\ \text{SYN} \begin{bmatrix} \text{CAT} & \textit{coord} \\ \text{SELECT} & \begin{bmatrix} \text{CRD} \ \textit{unmarked} \end{bmatrix} \\ \text{VAL} & \langle\rangle \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{FRAMES} \ \langle\rangle \end{bmatrix} \\ \text{CRD} \begin{bmatrix} \textit{marked} \\ \text{MODE} \quad \vee \end{bmatrix} \end{bmatrix}$$

The feature SELECT allows the HEAD-FUNCTOR CONSTRUCTION in (92) to combine a coordinator with a phrase, and to share the MODE and CRD values with the mother. This is illustrated in Figure 12.

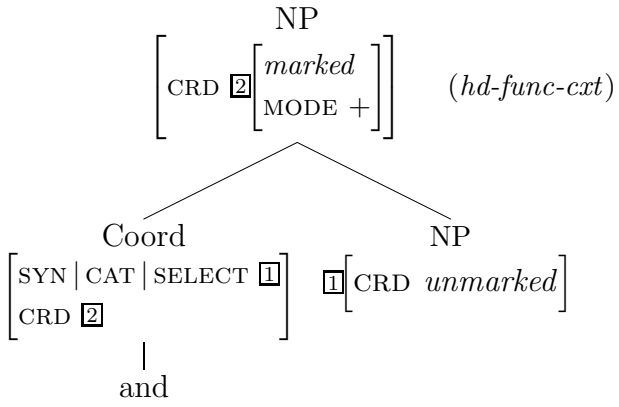

Figure 12: Coordinate head-functor construct

Following Sag et al. (2003), I assume that the semantic representation of any phrasal node is defined as the collection of the semantics of the daughters plus the semantic contribution of the construction itself, made via CX-REL. The rule behind the semantic composition (in any phrasal construction) is the SEMANTIC COMPOSITION RULE in (94).

(94) SEMANTIC COMPOSITION RULE

$$\textit{phrasal-cxt} \Rightarrow \begin{bmatrix} \text{MTR} \left[\text{SEM} \left[\text{FRAMES } \boxed{1} \oplus \ldots \oplus \boxed{n} \oplus \boxed{0}\right]\right] \\ \text{DTRS} \left\langle \left[\text{SEM} \left[\text{FRAMES } \boxed{1}\right]\right], \ldots, \left[\text{SEM} \left[\text{FRAMES } \boxed{n}\right]\right] \right\rangle \\ \text{CX-REL } \boxed{0} \end{bmatrix}$$

I assume that the HEAD FEATURE PRINCIPLE requires the mother's INDEX to be identical to the head daughter's INDEX, and that all the constructions discussed so far make no semantic contribution of their own (i.e. they are specified as CX-REL $\langle\rangle$). Coordination is different, however, as it will introduce a semantic relation between conjuncts.

Something that is common to all non-headed constructions is the fact that the category of the mother node reflects the categories of all of the daughters. In other words, the distribution of the whole is determined by the conjuncts, not by a distinguished head. Thus, NP coordination yields an NP, PP coordination yields a PP, S coordination yields an S, and so on.[19] Another common property is that such constructions do not have the ability to discharge gaps. As such, any gaps in the daughters of a non-headed construction must be percolated. I capture these two generalizations straightforwardly:

[19]This is similar to comparative constructions such as *some people bought* [*more* [*books*]$_{\text{NP}}$ *than* [*magazines*]$_{\text{NP}}$]$_{\text{NP}}$, *the frame is* [*more* [*wide*]$_{\text{AP}}$ *than* [*long*]$_{\text{AP}}$]$_{\text{AP}}$, or *students already pay* [*more* [*in fees*]$_{\text{PP}}$ *than* [*in tuition*]$_{\text{PP}}$]$_{\text{PP}}$.

(95) Non-Headed Construction

$$\textit{non-headed-cxt} \Rightarrow \begin{bmatrix} \text{MTR} \begin{bmatrix} \text{SYN} & \boxed{1} \end{bmatrix} \\ \text{DTRS} \left\langle \begin{bmatrix} \text{SYN} & \boxed{1} \\ \text{SLASH} & set(pg) \end{bmatrix}, \begin{bmatrix} \text{SYN} & \boxed{1} \\ \text{SLASH} & set(pg) \end{bmatrix} \right\rangle \end{bmatrix}$$

I am now in position to move on to the Coordination Construction rule, shown in (96). Along with Beavers and Sag (2004), I require the second conjunct to be marked by a coordinator, but not the second. This ensures that the grammar allows, for example, [*Tim* [*Sue* [*and Kim*]]] as well as [*Tim* [*and* [*Sue* [*and Kim*]]]] but not *[[*and Sue*] [*and Kim*]]. With regard to semantics, I propose that the coordination construction makes a type-underspecified meaning contribution. This contribution is determined by the meaning of the coordinator and the indices of the conjuncts. This way, the same coordination rule is general enough to deal with different coordination meanings.

(96) Coordination Construction

$$\textit{coord-cxt} \Rightarrow \begin{bmatrix} \text{MTR} \begin{bmatrix} \text{SEM} \mid \text{INDEX} & \boxed{x} \\ \text{CRD} \mid \text{MODE} & \boxed{2} \end{bmatrix} \\ \text{DTRS} \left\langle \begin{bmatrix} \text{SEM} \mid \text{INDEX} & \boxed{y} \\ \text{CRD} & \textit{unmarked} \end{bmatrix}, \begin{bmatrix} \text{SEM} \mid \text{INDEX} & \boxed{z} \\ \text{CRD} & \begin{bmatrix} \textit{marked} \\ \text{MODE} \ \boxed{2} \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{CX-REL} \left\langle \begin{bmatrix} \text{RELN} & \boxed{2} \\ \text{INDEX} & \boxed{x} \\ \text{ARG}_1 & \boxed{y} \\ \text{ARG}_2 & \boxed{z} \end{bmatrix} \right\rangle \end{bmatrix}$$

The mode value is structure-shared with the reln value in cx-rel, which is the semantic contribution made by the coordination. That contribution is a relation $\boxed{2}$ holding between the conjunct indices $\boxed{y}$ and $\boxed{z}$. Thus, if the value of mode is plurality formation ('+') then we obtain [reln +] in cx-rel. The latter means that the index $\boxed{x}$ is a sum of indices, as seen in (97a). If the coordinator is *or* and mode is a disjunction '$\vee$', then we have [reln $\vee$], and the relation in cx-rel states that the value of x can be y or z, as in (97b).

(97) a. $\begin{bmatrix} \text{RELN} & + \\ \text{INDEX} & \boxed{x} \\ \text{ARG}_1 & \boxed{y} \\ \text{ARG}_2 & \boxed{z} \end{bmatrix} \equiv x = (y + z)$ b. $\begin{bmatrix} \text{RELN} & \vee \\ \text{INDEX} & \boxed{x} \\ \text{ARG}_1 & \boxed{y} \\ \text{ARG}_2 & \boxed{z} \end{bmatrix} \equiv (x = y) \vee (x = z)$

The mother node and the second conjunct in (96) share the same mode values, and therefore all conjuncts will be semantically connected by the same coordination relation (plurality formation, disjunction, etc.), regardless of the number of applications of the rule. Since plurality-formation *and* is [mode +], then the value of reln is also

+, and conjunct indices are summed. Note that conjunction meanings like temporal precedence '$\prec$' and causality '$\rightarrow$' are only felicitous for event indices, correctly predicting that the conjuncts of such coordinations must be verbal rather than nominal. Thus, this account rules out temporal conjunction interpretations for [*Tom and Sue*] *arrived*. This account is illustrated in Figure 13.

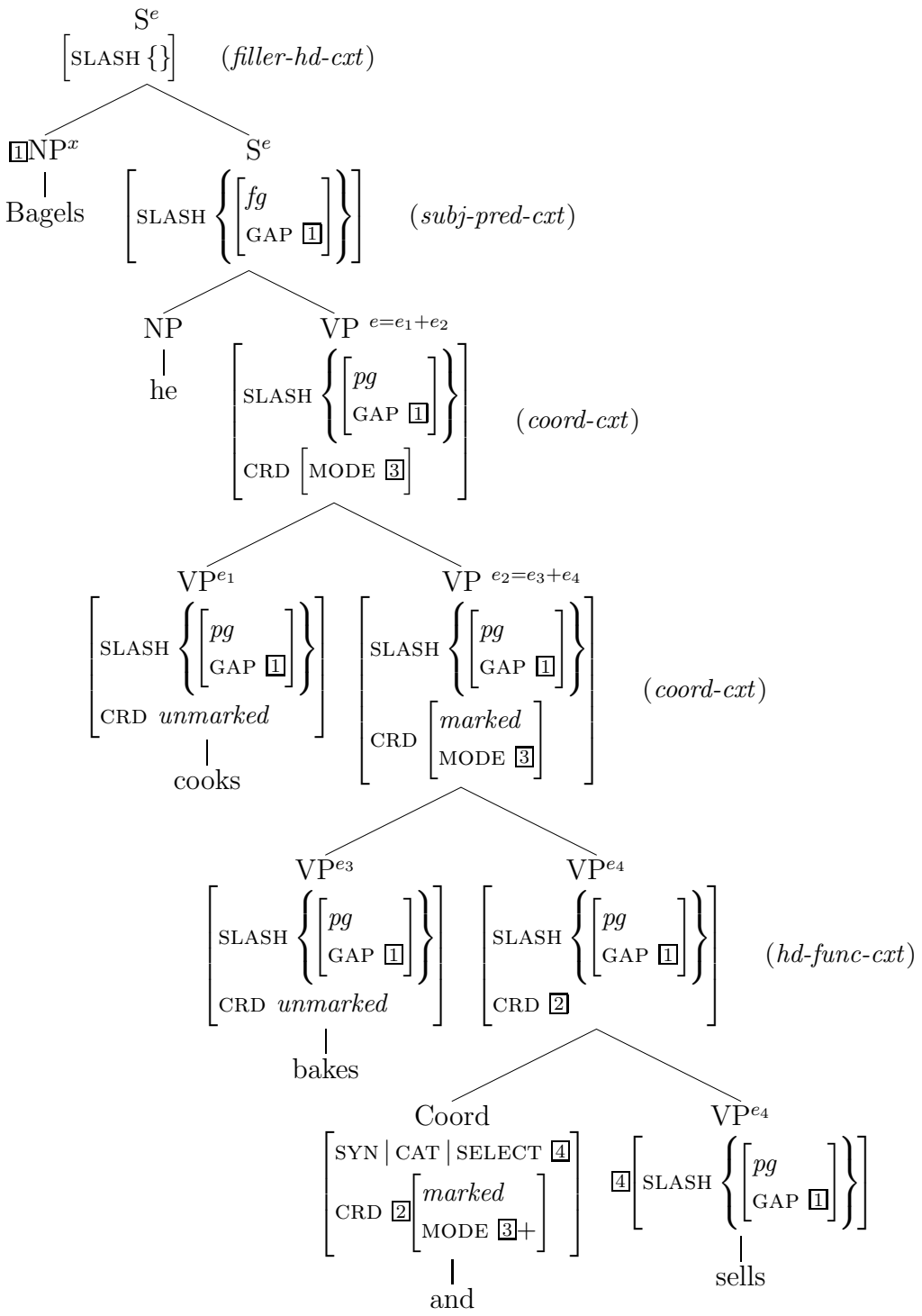

Figure 13: Symmetric conjunction: *Bagels*$_i$*, he cooks* $\__i$*, bakes* $\__i$*, and sells* $\__i$

The filler-head relevance condition R discussed in §4.3 obtains the extraction parallelism in symmetric coordination as follows. In a head-filler structure, the filler index x must be a coherent/relevant topic for the head structure that it has been extracted from. In the case of the sentence in Figure 13, the relation $R(x, e)$ must hold. Since $e = e_1 + e_2$, then $R(x, e_1 + e_2)$ must hold. Because this is an event sum, distributivity applies as usual $R(x, e_1) \wedge R(x, e_2)$, i.e. the filler must be relevant for both conjuncts. The former is satisfied (bagels are relevant for 'cooking') and the latter is satisfied if $R(x, e_3) \wedge R(x, e_4)$ holds. Although nothing in the GNIP forces gap fusion in symmetric coordination, this is the only possible option since R cannot be satisfied if the filler is not simultaneously predicated in all three conjuncts. Thus, distributivity causes the parallelism ATB effect. In asymmetric conjunction meanings there is no event plurality, and thus, no distribution forces the filler to be relevant for all conjuncts. Thus, extraction in asymmetric coordination need not be ATB.

6 Conclusion

This work argues that the syntax of unbounded dependencies in headed and non-headed constructions operates in exactly the same way. In principle, any given daughter – coordinate or not – can contain a gap, and multiple gaps can percolate and be optionally fused. Building on Pollard and Sag (1994) and Ginzburg and Sag (2000), this work proposes a uniform mechanism that governs the percolation, fusion, and discharge of filler-gap dependencies in coordinate and non-coordinate constructions alike. A generalized rule for coordinate structures is general enough to deal with different symmetric and asymmetric coordinations, their semantics, and their extraction patterns. Based on insights from Kuno (1987), Goldsmith (1985), and Kehler (2002), Ross's Coordinate Structure Constraint and ATB exceptions for symmetric coordination result not from syntactic stipulation but rather from the interplay between coherence and distributivity. The Conjunct Constraint is a consequence of both a traceless analysis and the fact that coordinators are viewed as markers rather than heads – a stance shown to have robust empirical support. Arguments are offered in support of the hypothesis held by Kluender (1992), Kluender and Kutas (1993), Kluender (1998), Levine and Hukari (2006), Hofmeister and Sag (2010) and others that gradient acceptability and certain islands are best viewed as resulting from measurable and independently motivated extragrammatical factors.

References

Abels, Klaus. 2003. Auxiliary adverb word order revisited. In *Proceedings of upenn colloquium on linguistics*, 26–115.

Abrusán, Márta. 2008. Contradiction and grammar: The case of weak islands. PhD dissertation, MIT.

Alexopoulou, Theodora and Frank Keller. 2007. Locality, cyclicity and resumption: At the interface between the grammar and the human sentence processor. *Language* 83 (1). 110–160.

Assmann, Anke, Fabian Heck, Johannes Hein, Stefan Keine, and Gereon Mueller. 2010. Does chain hybridization in Irish support movement-based approaches to long-distance

dependencies. In *Proceedings of the 17th international conference on head-driven phrase structure grammar, university of Parise-Sorbonne, France*. Stanford: CSLI Publications.
Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9: 5–16.
Baltin, Mark. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13.
Beavers, John, and Ivan A. Sag. 2004. Ellipsis and apparent non-constituent coordination. In Müller, Stefan (ed.). *Proceedings of the 11th international conference on Head-Driven Phrase Structure Grammar, Katholieke Universiteit Leuven*, 48–69. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/.
Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.
Borgonovo, Claudia, and Ad Neeleman. 2000. Transparent adjuncts. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 45: 199–224.
Borsley, Robert D. 2004. An approach to English comparative correlatives. In Müller, Stefan (ed.). *Proceedings of the 11th international conference on Head-Driven Phrase Structure Grammar, Katholieke Universiteit Leuven*, 70–925. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/.
Borsley, Robert D. 2005. Against ConjP. *Lingua* 4 (115): 461–482.
Bouma, Gosse, Rob Malouf, and Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 1 (19): 1–65.
Chao, Wyn. 1988. *On ellipsis*. New York and London: Garland.
Chaves, Rui P. 2008. Linearization-based word-part ellipsis. *Linguistics and Philosophy* 31 (3): 261–307.
Chaves, Rui P. 2009. Construction-based cumulation and adjunct extraction. In *Proceedings of the 16th international conference on head-driven phrase structure grammar, university of göttingen, germany*, 47–67. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2009/.
Chaves, Rui P. to appear. Conjunction, cumulation and respectively readings. *Journal of Linguistics*,
Chaves, Rui P., and Ivan A. Sag. 2007. Two kinds of ellipsis in English coordinate structures. *Unpublished ms*. University at Buffalo and Stanford University.
Chen, Evan, Edward Gibson, and Florian Wolf. 2005. Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language* 52: 144–169.
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: The MIT Press.
Chomsky, Noam. 1982. *Concepts and consequences of the theory of government and binding*. Cambridge: MIT Press.
Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. University of Chicago Press.
Chung, Sandra, and James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14: 703–714.
Cinque, Gugliemo. 1990. *Types of $\overline{A}$-dependencies*. Cambridge, MA: MIT Press.
Clements, George N. 1984. Binding domains in Kikuyu. *Studies in the Linguistic Sciences* 14: 37–56.
Cohen, Jonathan L. 1971. The logical particles of natural language. In *Pragmatics of natural language*, ed. Y. Bar-Hillel, 50–68. Dordrecht: Reidel.
Collins, Chris. 1994. Economy of derivation and the generalized proper binding condition. *Linguistic Inquiry* 25: 45–61.
Copestake, Ann, Alex Lascarides, and Dan Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proc. of the 39th ACL*. Toulouse, France.
Copestake, Ann, Daniel Flickinger, Ivan A. Sag, and Carl Pollard. 2006. Minimal recursion

semantics. *Journal Research on Language & Computation* 3 (4): 281–332.
Cowart, Wayne. 1997. *Experimental syntax: Applying objective methods to sentence judgments*. Sage.
Crain, Stephen, and Mark Steedman. 1985. On not being led up the garden path: the use of context by the psychological parser. In *Natural language parsing: Psychological, computational and theoretical perspectives, ACL studies in natural language processing, cambridge university press*, Dowty, David, Lauri Kartunnen and Arnold Zwicky (eds.) 320–358.
Culicover, Peter W., and Robert D. Levine. 2001. A reconsideration of locative inversion. *Natural Language and Linguistic Theory* 19: 283–310.
Deane, Paul D. 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2 (1): 1–63.
Deane, Paul D. 1992. *Grammar in mind and brain*. Mouton de Gruyter.
Diesing, Molly. 1990. Verb movement and the subject position in Yiddish. *Natural Language and Linguistic Theory* 8: 41–79.
Donohue, Mark, and Anna Maclachlan. 1999. What agreement in Chamorro? *Austroesian Formal Linguistics Association* 6: 121–132.
Eckardt, Regina. 1995. Does information flow in event structures? In *Ellipsis, underspecification, events and more in dynamic semantics, dyana-2 deliverable r2.2.c*, ed. Geroen Groenendijk, 49–72. ILLC/Department of Philosophy, University of Amsterdam.
Ellis, Rod. 1991. Grammaticality judgments and second language acquisition. *Studies in second language acquisition* 13(2): 161–186.
Engdahl, Elisabet. 1983. Parasitic gaps. *Linguistics and Philosophy* 6: 3–34.
Erteschik-Shir, Nomi. 1973. On the nature of island constraints. Doctoral dissertation, Cambridge, MIT.
Erteschik-Shir, Nomi. 1981. More on extractability from quasi-nps. *Linguistic Inquiry* 12: 665–670.
Fanselow, Gisbert, and Stefan Frisch. 2004. Effects of processing difficulty on judgments of acceptability. In *Gradience in grammar*, eds. G. Fanselow, C. Fery, M. Schlesewsky, and R. Vogel. OUP.
Fiengo, Robert, and James Higginbotham. 1981. Opacity in NP. *Linguistic Analysis* 7(4): 395–421.
Fodor, Janet. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9: 427–473.
Fodor, Janet D. 1983. Phrase structure parsing and the island constraints. *Linguistics and Philosophy* 6: 163–223.
Fodor, Janet D. 1992. Islands, learnability and the lexicon. In *Island constraints: Theory, acquisition and processing*, eds. Helen Good luck and Michael Rochemont. Dordrecht: Kluwer.
Frazier, Lynn. 1987. Syntactic processing: evidence from dutch. *Natural Language and Linguistic Theory* 5: 519–559.
Gawron, Jean M., and Andrew Kehler. 2003. Respective answers to coordinated questions. In *Semantics and linguistic theory 13 (salt 13)*. University of Washington, Seattle, Washington. University of Washington, Seattle, Washington.
Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12 (2): 155–184.
Gazdar, Gerald, Ewann Klein, Geoffrey K. Pullum, and Ivan A. Sag. 1985. *Generalized phrase structure grammar*. Oxford: Blackwell, and Cambridge, Ma: Harvard University Press.

Ginzburg, Jonathan, and Ivan A. Sag. 2000. *Interrogative Investigations: the form, meaning and use of English interrogative constructions*. Stanford: CSLI Publications.

Goldsmith, John. 1985. A principled exception to the Coordinate Structure Constraint. In *Papers from the twenty-first annual regional meeting of the chicago linguistic society*, 133–143. Chicago: Chicago Linguistic Society.

Goodall, Grant. 1987. *Parallel structures in syntax: Coordination, causatives, and restructuring*. New York: Cambridge University Press.

Grosu, Alexander. 1973. On the nonunitary nature of the coordinate structure constraint. *Linguistic Inquiry* 4: 88–92.

Hankamer, Jorge. 1974. Why there are two *than*'s in English. In *Papers from the ninth regional meeting of the chicago linguistic society*, 179–191. Chicago.

Hankamer, Jorge, and Ivan A. Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7 (3): 391–428.

Hegarty, Michael. 1990. On adjunct extraction from complements. In *Mit working papers in linguistics 13*, eds. L. L. S. Cheng and H. Demirdash, 101–124.

Hendriks, Petra. 1995. Comparatives and categorial grammar. Phd. dissertation, University of Groningen.

Hofmeister, Philip. 2007a. Facilitating memory retrieval in natural language comprehension. Phd thesis, Stanford University.

Hofmeister, Philip. 2007b. Retrievability and gradience in filler-gap dependencies. In *Proceedings from the annual meeting of the chicago linguistics society*, eds. M. Elliott, J. Kirby, O. Sawada, E. Staraki, and S. Yoon, Vol. 43, 1, 109–123.

Hofmeister, Philip, and Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86 (2): 366–415.

Huddleston, Rodney D., Geoffrey Pullum, and Peter Peterson. 2002. The cambridge grammar of the English language, 1033–1096. Cambridge University Press.

Hukari, Tom, and Robert D. Levine. 1995. Adjunct extraction. *Journal of linguistics* 31: 195–226.

Johannessen, Janne B. 1998. *Coordination*. New York: Oxford University Press.

Kathol, Andreas. 1993. Linearization and coordination in German. In *Papers in syntax, osu working papers in linguistics, number 42*, eds. Andreas Kathol and Carl J. Pollard, 109–123. Ohio State University: Department of Linguistics

Kathol, Andreas. 1995. Linearization-based German syntax. PhD. Thesis, Ohio State University..

Kathol, Andreas. 2000. *Linear syntax*. Oxford University Press.

Kayne, Richard. 1980. Extensions of binding and case-marking. *Linguistic Inquiry* 11: 75–96.

Kayne, Richard. 1983. Connectedness. *Linguistic Inquiry* 14: 223–249.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass..

Kayne, Richard, and Jean-Yves Pollock. 1978. Stylistic inversion, sucessive cyclicity, and Move NP in French. *Linguistic Inquiry* 9: 93–133.

Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI.

Kim, Jong-Bok, and Ivan A. Sag. 2005. English object extraposition: A constraint-based approach. In *Proceedings of the 12th international conference on head-driven phrase structure grammar, department of informatics, university of lisbon*, ed. Stefan Müller, 192–212. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/.

Kluender, Robert. 1992. Deriving island constraints from principles of predication. In *Island contraints: Theory, acquisition and processing*, eds. Helen Goodluck and Michael Rochemont, 223–258. Dordrecht: Kluwer.

Kluender, Robert. 1998. On the distinction between strong islands and weak islands: a

processing perspective. In *Syntax and semantics 29: The limits of syntax*, eds. P. W. Culicover and L. McNally, 241–279. New York, NY: Academic Press.

Kluender, Robert. 2004. Are subject islands subject to a processing account. In *In proceedings of the wccfl 23*, eds. B. Schmeiser, V. Chand, A Kelleher, and A. Rodriguez, 101–125. Somerville, MA: Cascadilla Press.

Kluender, Robert, and Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8: 573–633.

Konieczny, Lars. 1996. Human sentence processing: a semantics-oriented parsing approach. PhD diss, Universität Freiburg.

Kroch, Anthony. 1989. Amount quantification, referentiality, and long *wh*-movement. Ms., Department of Linguistics, University of Pennsylvania.

Kuno, Susumu. 1976. Subject, theme, and the speaker's empathy – a reexamination of relativization phenomena. In *Subject and topic*, ed. Charles N. Li, 417–444. New York: Academic Press.

Kuno, Susumu. 1987. *Functional syntax – anaphora, discourse and empathy*. The University of Chicago Press, Chicago and London.

Lakoff, George. 1986. Frame Semantic Control of the Coordinate Structure Constraint. In *Papers from the 22nd regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society.

Lasersohn, Peter. 1995. *Plurality, Conjunction and Events*. Kluwer: Dordrecht.

Levin, Nancy, and Ellen Prince. 1986. Gapping and clausal implicature. *Papers in Linguistics* 19: 351–364.

Levine, Robert, Thomas E. Hukari, and Michael Calcagno. 2001. Parasitic gaps in English: Some overlooked cases and their theoretical implications. In *Parasitic gaps*, eds. Peter Culicover and Paul M. Postal, 181–222. Cambridge, MA: MIT Press.

Levine, Robert D. 2001. The extraction riddle: just what are we missing? *Journal of Linguistics* 37: 145–174.

Levine, Robert D. 2003. Adjunct valents: cumulative scoping adverbial constructions and impossible descriptions. In *The proceedings of the 9th international conference on head-driven phrase structure grammar*, eds. Jongbok Kim and Stephen Wechsler, 209–232. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/3/.

Levine, Robert D., and Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions*. CSLI Publications, Leland Stanford Junior University.

Levine, Robert D., and Ivan A. Sag. 2003. Some empirical issues in the grammar of extraction. In *Proceedings of the hpsg-2003 conference, michigan state university, east lansing*, ed. Stefan Müller, 236–256. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/3/.

Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. MIT Press.

Link, Godehard. 1984. Hydras: on the logic of relative clause constructions with multiple heads. In *Varieties of formal semantics*, eds. F. Landman and F. Veltman, 245–257. Dordrecht: Foris. Dordrecht: Foris.

Link, Godehard. 1998. *Algebraic Semantics in Language and Philosophy*. CSLI Lecture Notes No. 74.

MacDonald, Maryellen C., Neal J. Pearlmutter, and Mark S. Seidenberg. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101 (4): 676–703.

Maling, Joan, and Annie Zaenen. 1978. The non-universality of a surface filter. *Linguistic Inquiry* 9: 475–497.

Maling, Joan, and Annie Zaenen. 1982. A phrase structure account of scandinavian extraction

phenomena. In Jacobson, Pauline & Goeffrey K. Pullum (eds.), *The nature of syntactic representation*, 229–282. Dordrecht: Reidel.

McCawley, James D. 1981. The syntax and semantics of english relative clauses. *Lingua* 53: 99–149.

McCloskey, James. 1979. *Model-theoretic semantics and transformational grammar*. Dordrecht: Reidel.

Munn, Alan B. 1993. *Topics in the syntax and semantics of coordinate structures*. PhD. dissertation, University of Maryland, Maryland.

Munn, Alan B. 1998. ATB movement without identity. In *Proceedings of the the 14th eastern states conference on linguistics (escol-97)*, eds. Jennifer Austin and Aaron Lawson, 150–160. CSC Publications.

Munn, Alan B. 1999. On the identity requirement of ATB extraction. *Natural Language Semantics* 7: 421–425.

Oehrle, Richard T. 1987. Boolean properties in the analysis of gapping. In *Syntax and semantics 20: Discontinuous constituency*, eds. Geoffrey J. Huck and Almerindo E. Ojeda, 201–240. Academic Press, Orlando, FL.

Oshima, David Y. 2007. On factive islands: pragmatic anomaly vs. pragmatic infelicity. In *Proceedings of the 20th annual conference on new frontiers in artificial intelligence. Jsai'06*, 147–161. Berlin, Heidelberg: Springer.

Pearlmutter, Neal J., and Maryellen C. MacDonald. 1979. Plausibility and syntactic ambiguity resolution. In *Proceedings of the 14th annual conference of the cognitive science society*, 498–503.

Pesetsky, David. 1982. Paths and categories. PhD. Dissertation, MIT, Cmbridge Mass..

Pesetsky, David. 1987. Wh-in-situ: movement and unselective binding. In *The representation of (in)definiteness*, eds. Eric Reuland and Alice ter Meulen, 98–129. MIT Press.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge: MIT Press.

Pickering, Martin, and Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* 6: 229–259.

Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press and Stanford: CSLI.

Postal, Paul M. 1993. Parasitic Gaps and the Across-the-board Phenomenon. *Linguistic Inquiry* 24: 735–754.

Postal, Paul M. 1994. Parasitic and pseudoparasitic gaps. *Linguistic Inquiry* 25: 63–117.

Postal, Paul M. 1998. *Three investigations of extraction*. MIT: Cambridge, Massachusetts.

Postal, Paul M. 2001. Islands. In *The handbook of syntactic theory*, eds. M. Baltin and C. Collins. Oxford: Blackwell.

Powell, Arnold, and Ray Gary Peters. 1973. Semantic clues in comprehension of novel sentences. *Language and Cognitive Processes* 6: 229–259.

Pritchett, Bradley L. 1991. Subjacency in a principle-based parser. In *Principle-based parsing: Computation and psycholinguistics*, ed. Robert C. Berwick, 301–345. Dordrecht: Kluwer.

Pustejovsky, James. 1995. *The generative lexicon*. MIT Press, Cambridge, MA, USA.

Ross, John. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, Massachusetts. [Published in 1986 as *Infinite Syntax!* Norwood, NJ: Ablex Publishing].

Sag, Ivan A. 2000. Another argument against *wh*-trace. Jorge Hankamer Webfest. http://ling.ucsc.edu/Jorge/sag.html.

Sag, Ivan A. 2002. Coordination and underspecification. In *Proceedings of the 9th international conference on head-driven phrase structure grammar,* Kyung Hee University, Seoul, South Korea, 267–2913.

Sag, Ivan A. 2005. Adverb extraction and coordination: a reply to Levine. In *The proceedings*

of the 12th international conference on head-driven phrase structure grammar, department of informatics, university of lisbon, ed. Stefan Müller, 322–342. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/.

Sag, Ivan A. 2010a. English filler-gap constructions. *Language* 86 (3): 486–545.

Sag, Ivan A. 2010b. Sign-based construction grammar: an informal synopsis. to appear in H. Boas and I. A. Sag (eds.), Sign-Based Construction Grammar. Stanford: CSLI Publications.

Sag, Ivan A., and Janet Fodor. 1994. Extraction without traces. In *Proceedings of the thirteenth west coast conference on formal linguistics*, 365–384. Stanford: CSLI.

Sag, Ivan A., and Thomas Wasow. 2011. Performance-compatible competence grammar. In *Non-transformational syntax*, eds. Robert Borsley and Kersti Borjars. Cambridge: Blackwells.

Sag, Ivan A., Philip Hofmeister, and Neal Snider. 2007. Processing complexity in subjacency violations: The complex noun phrase constraint. In *Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society. Chicago: CLS*.

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic theory - a formal introduction. 2nd edition*. Stanford: CSLI Publications.

Sag, Ivan A., Geral Gazdar, Thomas Wasow, and Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3: 117–171.

Sato, Yo, and Wai Lok Tam. 2008. Towards a unified account of adjuncts. In *The proceedings of the 15th international conference on head-driven phrase structure grammar*, ed. Stefan Müller, 173–188. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/9/.

Schmerling, Susan. 1972. Apparent counterexamples to the coordinate structure constraint: A canonical constraint. *Studies in the Linguistic Sciences* 2 (1): 91–104.

Steedman, Mark. 1989. Constituency and coordination in combinatory grammar. In *Alternative conceptions of phrase structure*, eds. Mark Baltin and Anthony Kroch. Chicago: University of Chicago Press.

Stolz, Walter S. 1976. A study of the ability to decode grammatically novel sentences. *Journal of Verbal Learning and Verbal Behavior* 6: 867–873.

Stowe, Lauriea. 1986. Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1: 227–245.

Van Eynde, Frank. 2003. On the notion 'determiner'. In *Proceedings of the hpsg-2003 conference, michigan state university, east lansing*, ed. Stefan Müller, 391–396. Stanford: CSLI Publications. http://csli-publications.stanford.edu/HPSG/3/.

Vasishth, Shravan, and Richard L. Lewis. 2006. Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language* 82(4): 767–794.

Warren, Tessa, and Edward Gibson. 2002. The influence of referential processing on sentence complexity. *Cognition* 85: 79–112.

Warren, Tessa, and Edward Gibson. 2005. Effects of np type and in reading cleft sentences in english. *Language and Cognitive Processes* 20: 751–767.

Williams, Edwin. 1978. Across-the-board rule application. *Linguistic Inquiry* 19: 31–43.

Wilson, Deirdre M. 1975. *Presupposition and non-truth-conditional semantics*. London: Academic Press.

Yngve, Victor. 1960. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society* 104: 444–466.

Zwart, Jan-Wouter. 2005. Some notes on coordination in head-final languages. In *Linguistics in the netherlands 2005*, eds. Jenny Doetjes and Jeroen van de Weijer, 232–241. John Benjamins, Amsterdam.